

ISSN 2078-9181

DOI 10.15622/sp.2019.18.5

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Отделение нанотехнологий и информационных технологий

САНКТ-ПЕТЕРБУРГСКИЙ
ИНСТИТУТ ИНФОРМАТИКИ И АВТОМАТИЗАЦИИ РАН

ТРУДЫ СПИИРАН

proceedings.spiiras.nw.ru



ТОМ 18 № 5



Санкт-Петербург
2019

18+

SPIIRAS PROCEEDINGS

Volume 18 № 5, 2019

Scientific, educational, and interdisciplinary journal primarily specialized
in computer science, automation, and applied mathematics

Trudy SPIIRAN ♦ Founded in 2002 ♦ Труды СПИИРАН

Founder and Publisher

St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences

Editor-in-Chief

R. M. Yusupov, Prof., Dr. Sci., Corr. Member of RAS, St. Petersburg, Russia

Editorial Board Members

A. A. Ashimov ,	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
N. P. Veselkin ,	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
O. Yu. Gusikhin ,	Ph. D., Dearborn, USA
V. Delic ,	Prof., Dr. Sci., Novi Sad, Serbia
A. Dolgui ,	Prof., Dr. Habil., St. Etienne, France
M. Zelezny ,	Assoc. Prof., Ph.D., Plzen, Czech Republic
I. A. Kalyaev ,	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
A. A. Karpov ,	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
D. A. Ivanov ,	Prof., Dr. Habil., Berlin, Germany
K. P. Markov ,	Assoc. Prof., Ph.D., Aizu, Japan
Yu. A. Merkuriev ,	Prof., Dr. Habil., Academician of the Latvian Academy of Sciences, Riga, Latvia
R. V. Meshcheryakov ,	Prof., Dr. Sci., Tomsk, Russia
N. A. Moldovian ,	Prof., Dr. Sci., St. Petersburg, Russia
V. E. Pavlovskiy ,	Prof., Dr. Sci., Moscow, Russia
A. A. Petrovsky ,	Prof., Dr. Sci., Minsk, Belarus
V. A. Putilov ,	Prof., Dr. Sci., Apatity, Russia
V. K. Pshikhopov ,	Prof., Dr. Sci., Taganrog, Russia
A. L. Ronzhin	(Deputy Editor-in-Chief), Prof., Dr. Sci., St. Petersburg, Russia
A. I. Rudskoi ,	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
H. Samani ,	Assoc. Prof., Ph.D., New Taipei City, Taiwan, Province of China
V. Sgurev ,	Prof., Dr. Sci., Academician of the Bulgarian academy of sciences, Sofia, Bulgaria
V. Skormin ,	Prof., Ph.D., Binghamton, USA
A. V. Smirnov ,	Prof., Dr. Sci., St. Petersburg, Russia
B. Ya. Sovetov ,	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
V. A. Soyfer ,	Prof., Dr. Sci., Academician of RAS, Samara, Russia
B. V. Sokolov ,	Prof., Dr. Sci., St. Petersburg, Russia
L. V. Utkin ,	Prof., Dr. Sci., St. Petersburg, Russia
A. L. Fradkov ,	Prof., Dr. Sci., St. Petersburg, Russia
H. Kaya ,	Assoc. Prof., Ph.D., Tekirdag, Turkey
L. B. Sheremetov ,	Assoc. Prof., Dr. Sci., Mexico, Mexico

Editor: A. I. Motienko

Editor: E. P. Miroshnikova

Technical editor: M. S. Avstriyskaya

Translator: N. V. Kashina

Editorial Board's address

14-th line VO, 39, SPIIRAS, St. Petersburg, 199178, Russia,
e-mail: publ@ias.spb.su, web: <http://www.proceedings.spiiras.nw.ru/>

The journal is indexed in Scopus

© St. Petersburg Institute for Informatics and Automation
of the Russian Academy of Sciences, 2019

ТРУДЫ СПИИРАН

Том 18 № 5, 2019

Научный, научно-образовательный, междисциплинарный журнал с базовой специализацией в области информатики, автоматизации и прикладной математики
Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
Санкт-Петербургский институт информатики и автоматизации Российской академии наук
(СПИИРАН)

Главный редактор

Р. М. Юсупов, чл.-корр. РАН, д-р техн. наук, проф., С-Петербург, РФ

Редакционная коллегия

- А. А. Ашимов**, академик национальной академии наук Республики Казахстан д-р техн. наук, проф., Алматы, Казахстан
Н. П. Веселкин, академик РАН, д-р мед. наук, проф., С.-Петербург, РФ
О. Ю. Гусихин, Ph.D., Диаборн, США
В. Делич, д-р техн. наук, проф., Нови-Сад, Сербия
А. Б. Долгий, Dr. Habil., проф., Сент-Этьен, Франция
М. Железны, Ph.D., доцент, Пльзень, Чешская республика
Д. А. Иванов, д-р экон. наук, проф., Берлин, Германия
И. А. Каляев, академик РАН, д-р техн. наук, профессор, Таганрог, РФ
А. А. Карпов, д-р техн. наук, доцент, С.-Петербург, РФ
К. П. Марков, Ph.D., доцент, Аизу, Япония
Ю. А. Меркурьев, академик Латвийской академии наук, Dr. Habil., проф., Рига, Латвия
Р. В. Мещеряков, д-р техн. наук, профессор, Томск, РФ
Н. А. Молдовян, д-р техн. наук, проф., С.-Петербург, РФ
В. Е. Павловский, д-р физ.-мат. наук, профессор, Москва, РФ
А. А. Петровский, д-р техн. наук, проф., Минск, Беларусь
В. А. Путилов, д-р техн. наук, проф., Апатиты, РФ
В. Х. Пшихопов, д-р техн. наук, профессор, Таганрог, РФ
А. Л. Ронжин (зам. главного редактора), д-р техн. наук, проф., С.-Петербург, РФ
А. И. Рудской, академик РАН, д-р техн. наук, проф., С.-Петербург, РФ
Х. Самани, Ph.D., доцент, Синьбэй, Тайвань, КНР
В. Сгурев, академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
В. А. Скормин, Ph.D., проф., Бингемптон, США
А. В. Смирнов, д-р техн. наук, проф., С.-Петербург, РФ
Б. Я. Советов, академик РАО, д-р техн. наук, проф., С.-Петербург, РФ
В. А. Сойфер, академик РАН, д-р техн. наук, проф., Самара, РФ
Б. В. Соколов, д-р техн. наук, проф., С.-Петербург, РФ
Л. В. Уткин, д-р техн. наук, проф., С.-Петербург, РФ
А. Л. Фрадков, д-р техн. наук, проф., С.-Петербург, РФ
Х. Кайя, Ph.D., доцент, Текирдаг, Турция
Л. Б. Шереметов, д-р техн. наук, Мехико, Мексика

Редактор: А. И. Мотиенко

Литературный редактор: Е. П. Миросникова

Технический редактор: М. С. Австрийская

Переводчик: Н. В. Кашина

Адрес редакции

199178, Санкт-Петербург, 14-я линия, д. 39,
e-mail: publ@iias.spb.su, сайт: <http://www.proceedings.spiiras.nw.ru/>

Журнал индексируется в международной базе данных Scopus

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»

© Федеральное государственное бюджетное учреждение науки

Санкт-Петербургский институт информатики и автоматизации Российской академии наук, 2019
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе печатного периодического издания–журнала «Труды СПИИРАН» статей по текущим экономическим, политическим, социальным и религиозным вопросам с обязательным указанием имени автора статьи и печатного периодического издания–журнала «Труды СПИИРАН»

CONTENTS

Mathematical Modeling Numerical Methods

S.N. Chukanov
COMPARISON OF OBJECTS' IMAGES BASED ON COMPUTATIONAL TOPOLOGY METHODS 1043

Y.I. Senkevich, Y.V. Marapulets, O.O. Lukovenkova, A.A. Solodchuk
TECHNIQUE OF INFORMATIVE FEATURES SELECTION IN GEOACOUSTIC EMISSION SIGNALS 1066

K.A. Batenkov
ACCURATE AND BOUNDARY ESTIMATE OF COMMUNICATION NETWORK CONNECTIVITY PROBABILITY BASED ON MODEL STATE COMPLETE ENUMERATION METHOD 1093

J.V. Doronina, A.V. Skatkov
STATISTICAL STABILITY ANALYSIS OF STATIONARY MARKOV MODELS 1119

Information Security

F.A. Taubin, A.N. Trofimov
CONCATENATED CODING FOR MULTILEVEL FLASH MEMORY WITH LOW ERROR CORRECTION CAPABILITIES IN OUTER STAGE 1149

A.V. Fedorchenko, E.V. Doynikova, I.V. Kotenko
AUTOMATED DETECTION OF ASSETS AND CALCULATION OF THEIR CRITICALITY FOR THE ANALYSIS OF INFORMATION SYSTEM SECURITY 1182

Digital Information Telecommunication Technologies

I.V. Prisukhina, D.V. Borisenko, S.A. Lunev
SIMULATION MODEL OF ELECTRIC CODE-MODULATED SIGNAL IN RUSSIAN SYSTEMS OF INTERVAL CONTROL OF TRAIN MOVEMENT BASED ON TRACK CIRCUIT 1212

V.I. Mironov, Y.V. Mironov, D.K. Khegay
OPTIMAL DETERMINATION OF SPACE OBJECTS ORBIT BY ANGULAR MEASUREMENTS OF GROUND-BASED OPTOELECTRONIC STATIONS 1239

СОДЕРЖАНИЕ

Математическое моделирование и прикладная математика

- С.Н. Чуканов
СРАВНЕНИЕ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ МЕТОДАМИ ВЫЧИСЛИТЕЛЬНОЙ
ТОПОЛОГИИ. 1043
- Ю.И. Сенкевич, Ю.В. Марапулец, О.О. Луковенкова, А.А. Солодчук
МЕТОДИКА ВЫДЕЛЕНИЯ ИНФОРМАТИВНЫХ ПРИЗНАКОВ В СИГНАЛАХ
ГЕОАКУСТИЧЕСКОЙ ЭМИССИИ 1066
- К.А. Батенков
ТОЧНЫЕ И ГРАНИЧНЫЕ ОЦЕНКИ ВЕРОЯТНОСТЕЙ СВЯЗНОСТИ СЕТЕЙ СВЯЗИ НА
ОСНОВЕ МЕТОДА ПОЛНОГО ПЕРЕБОРА ТИПОВЫХ СОСТОЯНИЙ 1093
- Ю.В. Дороница, А.В. Скатков
АНАЛИЗ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ СТАЦИОНАРНЫХ МАРКОВСКИХ
МОДЕЛЕЙ 1119

Информационная безопасность

- Ф.А. Таубин, А.Н. Трофимов
КАСКАДНОЕ КОДИРОВАНИЕ ДЛЯ МНОГОУРОВНЕВОЙ ФЛЭШ-ПАМЯТИ С
ИСПРАВЛЕНИЕМ ОШИБОК МАЛОЙ КРАТНОСТИ ВО ВНЕШНЕЙ СТУПЕНИ 1149
- А.В. Федорченко, Е.В. Дойникова, И.В. Котенко
АВТОМАТИЗИРОВАННОЕ ОПРЕДЕЛЕНИЕ АКТИВОВ И ОЦЕНКА ИХ КРИТИЧНОСТИ
ДЛЯ АНАЛИЗА ЗАЩИЩЕННОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ 1182

Цифровые информационно-телекоммуникационные технологии

- И.В. Присухина, Д. В. Борисенко, С. А. Лунёв
ИМИТАЦИОННАЯ МОДЕЛЬ ЭЛЕКТРИЧЕСКОГО КОДОВОГО СИГНАЛА В
РОССИЙСКИХ СИСТЕМАХ ИНТЕРВАЛЬНОГО РЕГУЛИРОВАНИЯ ДВИЖЕНИЯ
ПОЕЗДОВ НА ОСНОВЕ РЕЛЬСОВЫХ ЦЕПЕЙ 1212
- В.И. Миронов, Ю.В. Миронов, Д.К. Хегай
ОПТИМАЛЬНОЕ ОПРЕДЕЛЕНИЕ ОРБИТЫ КОСМИЧЕСКИХ ОБЪЕКТОВ ПО
УГЛОВЫМ ИЗМЕРЕНИЯМ НАЗЕМНЫХ ОПТИКО-ЭЛЕКТРОННЫХ СТАНЦИЙ 1239

С.Н. ЧУКАНОВ
**СРАВНЕНИЕ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ МЕТОДАМИ
ВЫЧИСЛИТЕЛЬНОЙ ТОПОЛОГИИ**

Чуканов С.Н. Сравнение изображений объектов методами вычислительной топологии.

Аннотация. Используя алгоритмы вычислительной топологии, рассмотрены методы сравнения изображений объектов, которые представлены множествами точек. Предложены алгоритмы построения множеств вещественных баркодов для сравнения изображений объектов. Нахождение баркодов форм объектов позволяет изучать как непрерывные, так и дискретные структуры, что делает их использование полезным в вычислительной топологии. Отличительной особенностью использования предложенных методов сравнения по отношению к методам алгебраической топологии является получение большего количества информации о форме объекта. Важным направлением применения вещественных баркодов является изучение инвариантов больших объемов данных. Предлагается метод, который объединяет технологии построения баркодов с внедренной негеометрической информацией (цвет, момент времени формирования, давление пера), представленной в виде функций от симплицидальных комплексов. Для этого баркоды расширяются функциями от симплексов для представления разнородной информации. Представленная структура расширенных баркодов повышает эффективность методов персистентных гомологий при сравнении изображений и распознавании образов. Предложена модификация метода Вассерштейна для нахождения расстояния между изображениями путем введения негеометрической информации о расстояниях между изображениями, обусловленной неравенствами функций исходного и терминального изображений соответствующих симплексов. Геометрические характеристики объекта могут изменяться при диффеоморфных деформациях; предложенные алгоритмы формирования расширенных баркодов изображений являются инвариантными к преобразованиям вращения и переноса. Рассмотрен метод определения расстояния между множествами точек, представляющими кривые, с учетом ориентации отрезков кривых. В работе используются понятия алгебраической и вычислительной топологии, теории лиевых групп, диффеоморфных преобразований.

Ключевые слова: распознавание образов, диффеоморфные преобразования, персистентные гомологии, расстояние Вассерштейна.

1. Введение. Существуют различные методы решения задачи сравнения форм исходного изображения и терминального (целевого) изображения [1-3]. Для анализа деформаций изображения от исходного к целевому в [4, 5] рассматривается группа диффеоморфизмов, а деформация рассматривается как геодезическая (кратчайший путь) между исходным и целевым изображениями. Задача сравнения исходного и целевого изображений решается на основе метода построения минимизируемого функционала [4, 5], характеризующего эволюцию диффеоморфного преобразования изображения от начального до терминального, и штрафа за отклонение траектории изображения от требуемой. Для решения задачи сравнения применяется метод частиц-ориентиров [6].

В случае анализа формы объекта используются геометрические, спектральные и топологические характеристики объекта. Геометрические характеристики объекта могут изменяться при диффеоморфных деформациях. Характеристики формы, определяемые топологическими методами, не зависят от координатного представления рассматриваемой формы и инвариантны по отношению к диффеоморфным преобразованиям. Основные методы топологии заключаются в создании сжатых комбинаторных представлений форм (например, триангуляции) и измерении формы объекта с помощью гомологических сигнатур.

В работе применяются алгоритмы построения множеств персистентных баркодов [7, 8] для сравнения диффеоморфных изображений, которые определяются вещественными значениями, принимаемыми персистентными гомологиями. Нахождение персистентных баркодов форм объектов позволяет изучать как непрерывные, так и дискретные структуры, что делает их использование полезным в вычислительной топологии. Отличительной особенностью использования персистентных гомологий по отношению к методам алгебраической топологии [9, 10] является получение большего количества информации о форме объекта. Важным направлением применения персистентных гомологий является изучение инвариантов больших объемов данных.

В дополнение к геометрической информации, существует широкий спектр негеометрической информации, которая ассоциируется с симплексами, представляющими изображение (цвет, момент времени формирования, давление пера и т.д.) и не описывается персистентной гомологией. Желательно иметь математическую основу для систематического внедрения как геометрической, так и негеометрической информации в единые топологические описания. Для этого предлагается метод, основанный на персистентных когомологиях, который объединяет технологии персистентных гомологий с внедренной негеометрической информацией, представленной в виде функций от симплицальных комплексов. Для это баркоды персистентных гомологий расширяются функциями от симплексов для представления разнородной информации. Предлагаемая структура расширенных баркодов с использованием когомологий повышает эффективность методов персистентных гомологий.

Для нахождения расстояния между изображениями при распознавании образов предлагается метод Леонида Вассерштейна модифицировать и в формулу для расстояния между изображениями, определяемого по баркодам персистентных гомологий, ввести негеометрическую информацию о расстоянии между изображениями,

обусловленную неравенствами функций исходного и терминального изображений соответствующих симплексов.

2. Гамильтонова механика точечных ориентиров изображения. Представление диффеоморфного отображения изображений можно рассмотреть как эволюцию точечных ориентиров (наиболее важных точек) изображения на основе гамильтоновой механики. Рассмотрим параметризацию формы изображения с помощью частиц — точечных ориентиров. Пусть $q_i(t); i=1, \dots, N$ используется для обозначения вектора положения i -й частицы и $p_i(t); i=1, \dots, N$ соответствующего вектора импульса в момент времени t . Если формально принять, что импульсы и скорости частиц связаны соотношением: $p_i = \mathcal{L} \cdot v_i$, где \mathcal{L} — линейный обратимый оператор (в механике соответствует инертной массе), то обратное соотношение: $v_i = \mathcal{L}^{-1} \cdot p_i = \mathcal{K} p_i$. Скалярное произведение векторов p_i и v_i , соответствующее метрике С.Л. Соболева [11], можно представить в виде: $\langle p_i, v_i \rangle_{L^2} = \langle \mathcal{L} v_i, v_i \rangle_{L^2}$. Для оператора: $\mathcal{L} = \text{id} - \alpha \nabla^2$ в пространстве $\Omega = \mathbb{R}^2$ — обратным оператором формально является оператор: $\mathcal{K} = \mathcal{L}^{-1}$, который можно аппроксимировать скалярной функцией Гаусса [3]:

$$K(q_i - q_j) = \beta e^{-\alpha^{-2}(q_i - q_j)^T (q_i - q_j)}.$$

Построим минимизируемый функционал, который соответствует деформации изображения, представленного совокупностью точечных ориентиров в виде:

$$J_0 = \frac{1}{2} \int_0^1 \left\{ \sum_{i,j=1}^N p_i^T K(q_i - q_j) p_j \right\} dt.$$

Минимизация J_0 должна осуществляться по значениям векторов q_i, q_j и p_i, p_j . Но значения векторов q_i, q_j определяются начальными значениями $q_i(0), q_j(0)$, поэтому минимизация осуществляется по начальным неизвестным значениям векторов $p_i(0), p_j(0)$. Задачу минимизации J_0 можно представить как задачу оптимального управления, связанную с гамильтонианом:

$$H_0(q, p) = \frac{1}{2} \sum_{i,j=1}^N p_i^T K(q_i - q_j) p_j.$$

Если гамильтониан системы принять в форме (с учетом J_0):

$$H(q, p) = H_0(q, p) + \sigma^{-2} \sum_{i=1}^N (\dot{q}_i - v_i(q))^2 = H_0(q, p) + \sigma^2 \sum_{i=1}^N p_i^2,$$

то уравнения Гамильтона для производных компонент $\dot{p} = (\dot{p}_1, \dots, \dot{p}_N)$, $\dot{q} = (\dot{q}_1, \dots, \dot{q}_N)$ примут вид:

$$\begin{aligned} \dot{p}_i &= -\frac{\partial H}{\partial q_i} = -\sum_{j=1}^N p_j^T \nabla_{q_i} K(q_i - q_j) p_j, \\ \dot{q}_i &= \frac{\partial H}{\partial p_i} = \sum_{j=1}^N K(q_i - q_j) p_j + \sigma^2 p_i. \end{aligned} \tag{1}$$

В функции Гамильтона $H(q, p)$ вводится штраф $\sigma^{-2} \sum_{i=1}^N (\dot{q}_i - v_i(q))^2$ при наличии рассогласования $(\dot{q}_i - v_i(q))$.

3. Расстояние между множествами точек. Пусть даны два множества точек, представляющие кривые, $Q_1, Q_2 \in \mathbb{R}^d$ (для плоских кривых $d=2$). Сформируем функцию Гаусса $K: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, $K(q_i, q_j) = \exp\left(-\left(q_i - q_j\right)^T \left(q_i - q_j\right) \alpha^{-2}\right)$ со свойством $K(q_i, q_i) = 1, \forall q_i \in (Q_1 \cup Q_2)$ и при увеличении $\left(q_i - q_j\right)^T \left(q_i - q_j\right)$ функция $K(q_i, q_j)$ уменьшается. Квадрат расстояния между множествами Q_1 и Q_2 равно [12]:

$$d^2(Q_1, Q_2) = \sum_{q_1 \in Q_1} \sum_{q_1' \in Q_1} K(q_1, q_1') + \sum_{q_2 \in Q_2} \sum_{q_2' \in Q_2} K(q_2, q_2') - 2 \sum_{q_1 \in Q_1} \sum_{q_2 \in Q_2} K(q_1, q_2).$$

Для учета ориентации отрезков кривых определим квадрат расстояния между двумя формами, моделируемыми с использованием касательных к точкам кривых. Квадрат расстояния между множествами $q_1^1, q_1^2, \dots, q_1^N \in Q_1$ и $q_2^1, q_2^2, \dots, q_2^N \in Q_2$ с учетом ориентации отрезков кривых можно определить из соотношения:

$$\begin{aligned}
 d^2 = & \sum_{q_1(i) \in Q_1} \sum_{q_1(j) \in Q_1} \eta_{q_1(i)}^T K(c_{q_1(i)}, c_{q_1(j)}) \eta_{q_1(j)} + \\
 & + \sum_{q_2(i) \in Q_2} \sum_{q_2(j) \in Q_2} \eta_{q_2(i)}^T K(c_{q_2(i)}, c_{q_2(j)}) \eta_{q_2(j)} - \\
 & - 2 \sum_{q_1(i) \in Q_1} - \sum_{q_2(j) \in Q_2} \eta_{q_1(i)}^T K(c_{q_1(i)}, c_{q_2(j)}) \eta_{q_2(j)},
 \end{aligned}$$

где $c_{q_k(i)} = (q_k(i) + q_k(i+1))/2$ — центры отрезков кривой между точками $q_k(i)$ и $q_k(i+1)$, $\eta_{q_k(i)} = (q_k(i+1) - q_k(i))/2$ — касательные векторы отрезков кривой в центрах $c_{q_k(i)}$, $k=1,2$, $i, j=1, \dots, N$.

Касательные векторы отрезков кривых учитывают ориентацию кривых. Недостатком данного метода нахождения квадрат расстояния между множествами точек, представляющими кривые, является отсутствие инвариантности по отношению к евклидовым преобразованиям координат точек.

4. Построение комплекса Вьеториса — Рипса. На основе оценивания функционала J_0 можно построить алгоритм распознавания изображений [11], но этот алгоритм требует проведение предварительной нормализации изображения и не является инвариантным по отношению преобразованиям координат. Поэтому алгоритмы распознавания, инвариантные по отношению преобразованиям координат (coordinate-free), являются актуальными [13, 14]; к таким алгоритмам относятся алгоритмы вычислительной топологии [15-17].

Для сравнения изображений, представленных множеством точечных ориентиров, предлагается построение комплекса Вьеториса — Рипса из точечных облаков этих ориентиров. Рассмотрим построение симплициальных комплексов из точечных облаков вложенных в \mathbb{R}^d . Пусть задано множество точечных ориентиров $S = \{q_1, \dots, q_k\}; q_i \in \mathbb{R}^2$. Комплекс Вьеториса — Рипса $R(r)$ представляет собой симплициальный комплекс, построенный на множестве точек S ; комплекс Вьеториса — Рипса имеет симплекс тогда, когда все пары точек находятся на расстоянии, не превышающем $2r$: $R(r) = \{\sigma = S \mid \|q_i - q_j\| \leq 2r, \forall i, j\}$.

При $r=0$ числа Бетти множества k точек равны: $\beta_0 = k; \beta_i = 0, i \geq 1$; характеристика Эйлера: $\chi = \beta_0$. При диффеоморфных преобразованиях множества k точек эти топологические характеристики не изменяются. При изменении $r \in \mathbb{R}^+$ значения чисел Бетти и характеристик Эйлера сформированного комплекса Вьеториса — Рипса изменяются при соответствующих диффеоморфных преобразованиях. Существуют точки появления r_b (birth) и точки исчезновения r_d (death) нетривиальных классов гомологий [15] и соответствующих чисел Бетти и характеристик Эйлера [9]. Горизонтальные полуинтервалы в 2D полуплоскости $\{[b_i \dots d_i] \mid d_i \geq b_i \geq 0\}$ называются баркодами (здесь i — номер баркода). Баркоды, числа Бетти и характеристики Эйлера для точечных облаков можно определить с помощью пакета JavaPlex [18].

Персистентная гомология сохраняет следы появления b_i и исчезновения d_i классов гомологии вдоль фильтрации симплициального комплекса, обусловленной сортировкой $\{x_1 < x_2 < \dots < x_l\}$, которая включает в себя информацию о гомологии каждого фиксированного симплициального комплекса при фильтрации $\{X(x_i)\}$. Группы гомологий $H_k(X(x_i))$ могут быть представлены в виде векторных пространств. Карта включения, связывающая группы, индуцирует последовательность линейных преобразований на векторных пространствах:

$$H_k(X(x_0)) \rightarrow H_k(X(x_1)) \rightarrow \dots \rightarrow H_k(X(x_l)).$$

Если существует функционал $\mathcal{Y}(\mathcal{X})$, который принимает значения на цепи \mathcal{X} , что можно записать в виде $\mathcal{Y}(\mathcal{X}) = \langle \mathcal{Y}, \mathcal{X} \rangle$, то определение кограничного оператора $d_n: C^{n-1} \rightarrow C^n$ принимает форму определения сопряженного линейного оператора:

$$\langle d_n \mathcal{Y}, \mathcal{X} \rangle = \langle \mathcal{Y}, \partial_n \mathcal{X} \rangle, \mathcal{X} \in C_n, \mathcal{Y} \in C^{n-1}; d_{n+1} d_n = 0.$$

Определим группы n -мерных коциклов: $Z^n = \text{Ker}(d_n)$; и n -мерных кограниц: $B^n = \text{Im}(d_{n-1})$. Факторгруппа $H^n = Z^n / B^n$

называется n -мерной группой когомологий коцепного комплекса. Для персистентных когомологий можно построить диаграмму [20, 21]:

$$H^k(X(x_0), \mathbb{R}) \leftarrow H^k(X(x_1), \mathbb{R}) \leftarrow \dots \leftarrow H^k(X(x_l), \mathbb{R}).$$

Из теоремы универсальности коэффициентов [9] для когомологий следует, что изоморфизм между группой гомологий $H_k(X)$ и группой когомологий $H^k(X, \mathbb{R})$. Это свойство означает, что персистентные гомологии и персистентные когомологии имеют одинаковые баркоды.

Пример 1. Приведем пример построения комплекса Вьеториса — Рипса для четырех облаков точек с центрами в вершинах $q_1 = (0,1), q_2 = (-1,0), q_3 = (0,-1), q_4 = (1,0)$ при $r = 1.0$, $r = 1.5$ и $r = 2.5$ с указанием чисел Бетти и характеристик Эйлера.

На рисунках 1a, 1b и 1c приведены точечные облака для случаев $r = 1.0$, $r = 1.5$ и $r = 2.5$ соответственно.

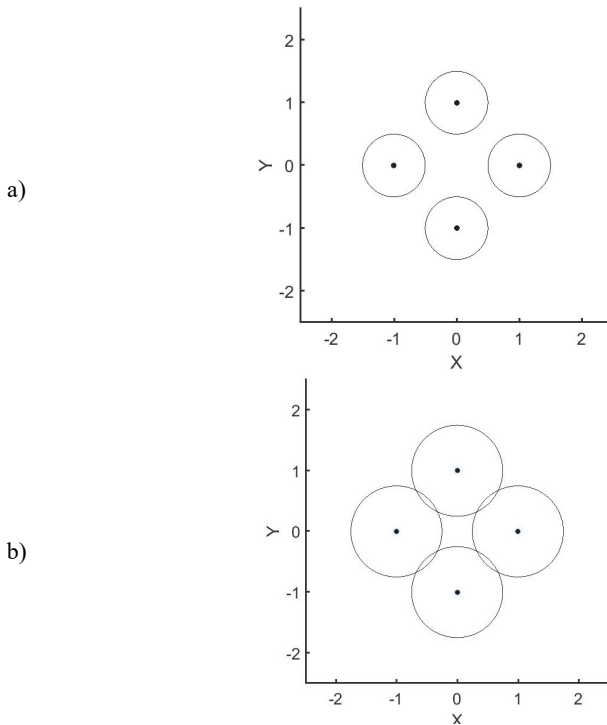


Рис. 1. Точечные облака для случая: а) $r = 1.0$; б) $r = 1.5$

(с)

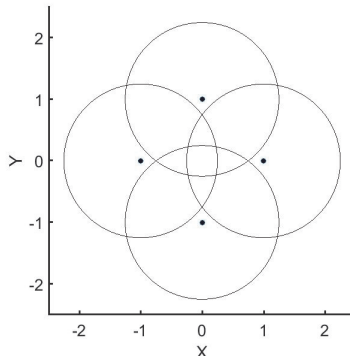


Рис. 1. Точечные облака для случая: с) $r = 2.5$

На рисунках 2а, 2б и 2с приведены сформированные комплексы Вьеториса — Рипса для случаев $r = 1.0$, $r = 1.5$ и $r = 2.5$ соответственно.

Для сформированного комплекса Вьеториса — Рипса можно определить числа Бетти β_0, β_1 и характеристики Эйлера χ :

- для случая $r = 1.0$: $\beta_0 = 4, \beta_1 = 0$; $\chi = 4$;
- для случая $r = 1.5$: $\beta_0 = 1, \beta_1 = 1$; $\chi = 0$;
- для случая $r = 2.5$: $\beta_0 = 1, \beta_1 = 0$; $\chi = 1$.

Для полученного комплекса Вьеториса — Рипса можно получить следующие баркоды: в размерности 0: $3[0 \dots 1.414]$; $[0 \dots \infty)$; в размерности 1: $[1.414 \dots 2)$. □

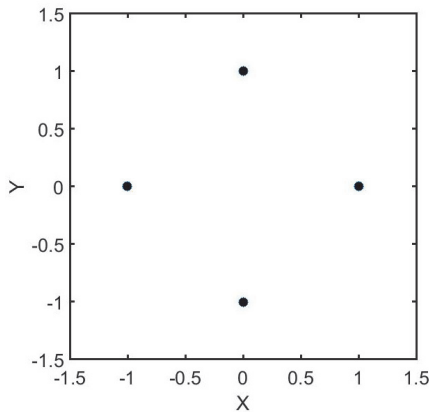
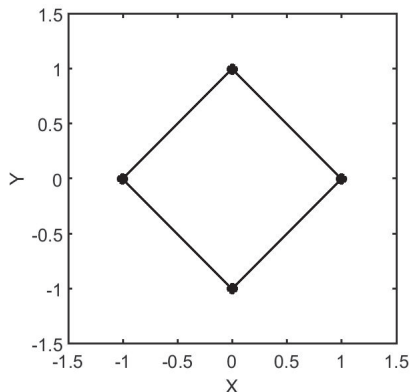
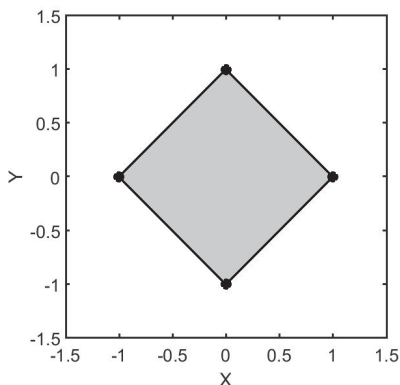


Рис. 2а. Комплекс Вьеториса — Рипса для случая $r = 1.0$


 Рис. 2b. Комплекс Вьеториса — Рипса для случая $r = 1.5$

 Рис. 2c. Комплекс Вьеториса — Рипса для случая $r = 2.5$

5. Модифицированное расстояние Вассерштейна. С учетом когомологий на симплексах (точках, отрезках, треугольниках ...) в работе предлагается расширить баркоды соответствующими значениями коцепей. Расширенный баркод может быть представлен тремя элементами: значением появления баркода b (birth), значением исчезновения баркода d (death) и значением функции на заданной цепи симплекса f . Для заданных расширенных баркодов одинаковой размерности: $B = \{\{b_i, d_i, f_i\}_{i \in I}\}$; $B' = \{\{b'_j, d'_j, f'_j\}_{j \in J}\}$, где $j \in J$ — номер симплекса, полученный отображением симплекса с номером $i \in I$: $j = \varphi(i)$, можно вычислить расстояние между B и B' , аналогичное

расстоянию Вассерштейна. Определяем расстояния между двумя соответствующими баркодами: $[b, d]$ и $[b', d']$:

$\Delta_b([b, d], [b', d']) = \max\{|b - b'|, |d - d'|\}$, и расстояние между функциями

f_i и f'_j : $\Delta_f(f, f') = \left| (d - b)^{-1} \int_b^d f(x) dx - (d' - b')^{-1} \int_{b'}^{d'} f'(x) dx \right|$. При

$f(x) = \text{const}; f'(x) = \text{const}$ можно записать:

$$\Delta_f(f_i, f'_j) = |f_i - f'_j|. \quad (2)$$

Определим расстояние между изображениями, обусловленное неравенствами границ баркодов исходного и терминального изображений (расстояние Вассерштейна):

$$d_{W,2}(B, B') = \left(\sum_{i \in I} \left(\Delta_b([b_i, d_i], [b'_{\varphi(i)}, d'_{\varphi(i)}]) \right)^2 \right)^{1/2}. \quad (3)$$

Расстояние между изображениями, обусловленное неравенствами функций исходного и терминального изображений соответствующих симплексов:

$$d_f(B, B') = \left(\sum_{i \in I} \left(\Delta_f(f_i, f'_{\varphi(i)}) \right)^2 \right)^{1/2}. \quad (4)$$

Модифицированное расстояние Вассерштейна можно определить из соотношения для евклидовой метрики:

$$d_{MW}(\mu_b, \mu_f, B, B') = \left(\mu_b \cdot d_{W,2}(B, B')^2 + \mu_f \cdot d_f(B, B')^2 \right)^{1/2}, \quad (5)$$

где μ_b — весовой параметр для $d_{W,2}(B, B')$ и μ_f — весовой параметр для $d_f(B, B')$.

Приведем пошаговый алгоритм определения расстояния Вассерштейна для сравнения симплицальных комплексов изображений объектов.

1. Задать координаты точек множества ориентиров $S = \{q_1, \dots, q_k\}$.

2. Найти матрицы взаимных расстояний между точками q_i и q_j , $i, j = 1 \dots k$.

3. На множестве S сформировать симплициальный комплекс Вьеториса — Рипса $R(r)$ для заданного значения $r \in \mathbb{R}$.

4. Определить точки появления r_b (birth) и точки исчезновения r_d (death) нетривиальных классов гомологий и соответствующих чисел Бетти и характеристик Эйлера.

5. Сформировать баркоды симплициального комплекса исходного изображения $B = \{\{b_i, d_i, f_i\}\}_{i \in I}$ и симплициального комплекса терминального изображения $B' = \{\{b'_j, d'_j, f'_j\}\}_{j \in J}$.

6. Определить расстояния между двумя соответствующими баркодами исходного и терминального изображений: $[b, d)$ и $[b', d')$:

$$\Delta_b([b, d), [b', d')) = \max\{|b - b'|, |d - d'|\}.$$

7. Определить расстояние Вассерштейна между изображениями по формуле (2).

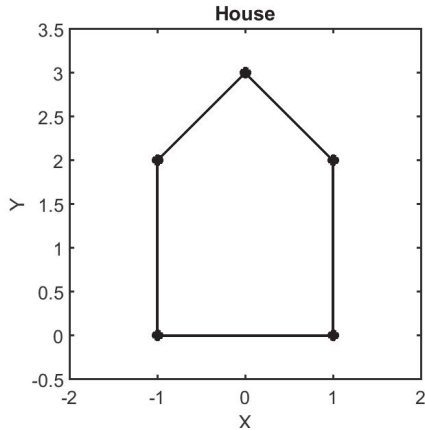
8. Для нахождения модифицированного расстояния Вассерштейна определить расстояние между изображениями, обусловленное неравенствами функций исходного и терминального изображений соответствующих симплексов по формуле (3).

9. Определить модифицированное расстояние Вассерштейна между изображениями по формуле (4).

Пример 2. Рассмотрим объект из 5 точек $\text{points_House} = [-1, 0; 1, 0; 1, 2; 0, 3; -1, 2]$ (в нотации JavaPlex), аналогичный примеру «House example» из [11] и объект из 5 точек $\text{points_House1} = [1, 3; -1, 3; -1, 1; 0, 0; 1, 1]$, полученный вращением относительно точки $[0 \ 0]$ на угол π с переносом на вектор $[0 \ 3]$. На рисунках 3а и 3б представлены изображения объектов House и House1.

На рисунках 4а и 4б представлены баркоды персистентных гомологий изображений объектов House и House1.

a)



b)

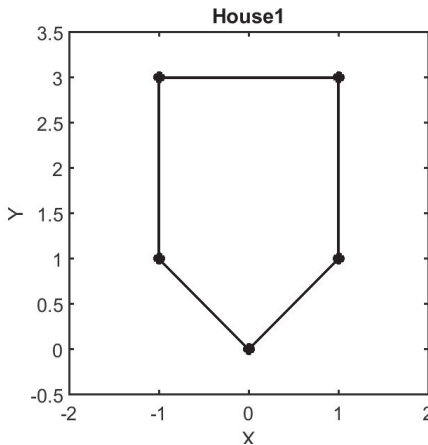


Рис. 3. Изображение объекта: а) House; б) House1

Расстояние между изображениями, обусловленное неравенствами границ баркодов исходного и терминального изображений House и House1: $d_{W,2}(B, B') = 0$, то есть гомологические характеристики объектов House и House1 совпадают.

Поставим в соответствие каждому симплексу нулевой размерности (точке) значение функции в соответствии с таблицей 1. Значение функции для объекта House возрастает при обходе против часовой стрелки; для объекта House1 — возрастает при обходе по часовой стрелке.

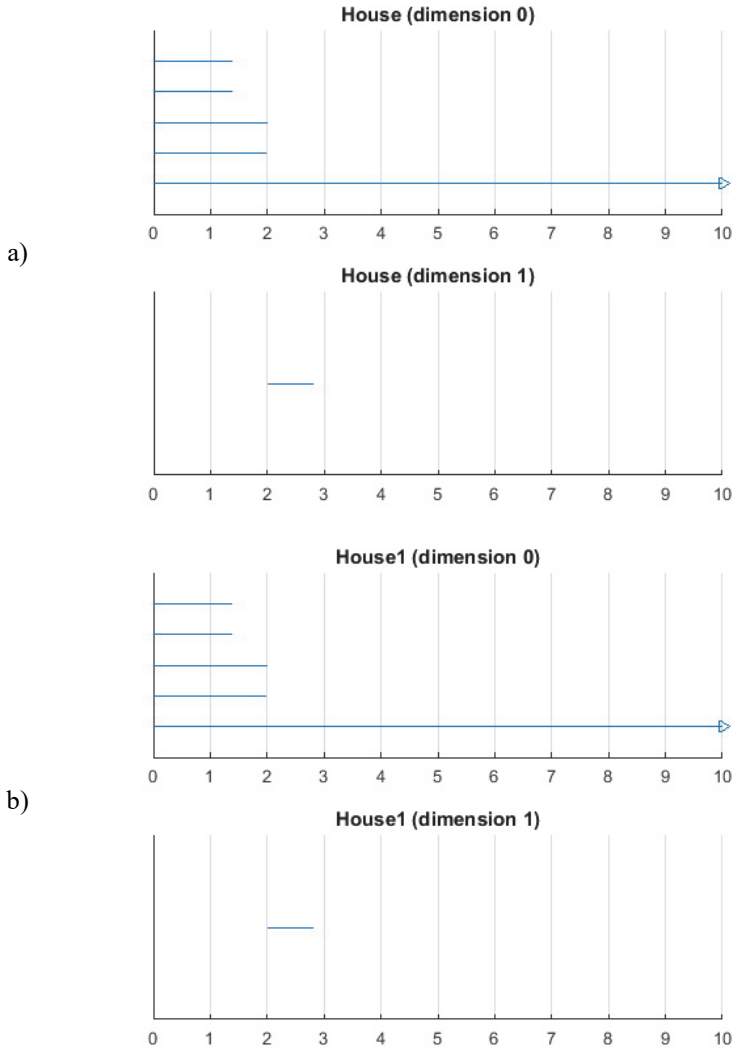


Рис. 4. Баркоды персистентных гомологий изображений объекта: а) House; б) House1

Модифицированное расстояние Вассерштейна между объектами House и House1 в этом случае равно расстоянию, обусловленному неравенствами функций исходного и терминального изображений:

$$d_{MW}(\mu_b, \mu_f, B, B') = d_f(B, B') = 4,47; \mu_b = 1, \mu_f = 1. \square$$

Таблица 1. Соответствие между точками и значениями функций

№ точки	Координаты точки \mathcal{X} объекта House	Значение функции f	Координаты точки $\varphi(\mathcal{X})$ объекта House1	Значение функции f'
1	$[-1 \ 0]^T$	$f(1)=1$	$[1 \ 3]^T$	1
2	$[1 \ 0]^T$	$f(2)=2$	$[-1 \ 3]^T$	5
3	$[1 \ 2]^T$	$f(3)=3$	$[-1 \ 1]^T$	4
4	$[0 \ 3]^T$	$f(4)=4$	$[0 \ 0]^T$	3
5	$[-1 \ 2]^T$	$f(5)=5$	$[1 \ 1]^T$	2

Пример 3. Рассмотрим пример сравнения подписей Наполеона, сформированных в 1809 и 1814 годах [11], на основе определения модифицированного расстояния Вассерштейна между ними при различных динамических темпах написания подписи. Сформируем объекты из 32 точек `points_1809` и `points_1814`.

На рисунках 5а и 5б представлены изображения подписей Наполеона 1809 и 1814 годов.

На рисунках 6а и 6б представлены баркоды персистентных гомологий изображений подписей Наполеона 1809 и 1814 годов, полученные с использованием пакета JavaPlex [18].

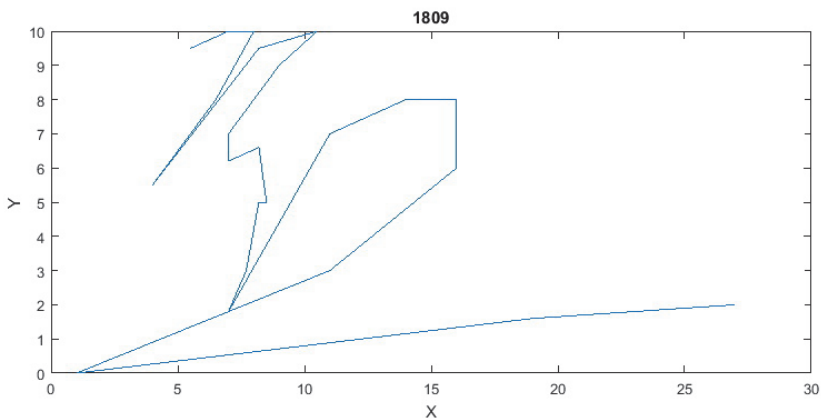


Рис. 5а. Изображение подписи Наполеона 1809 года

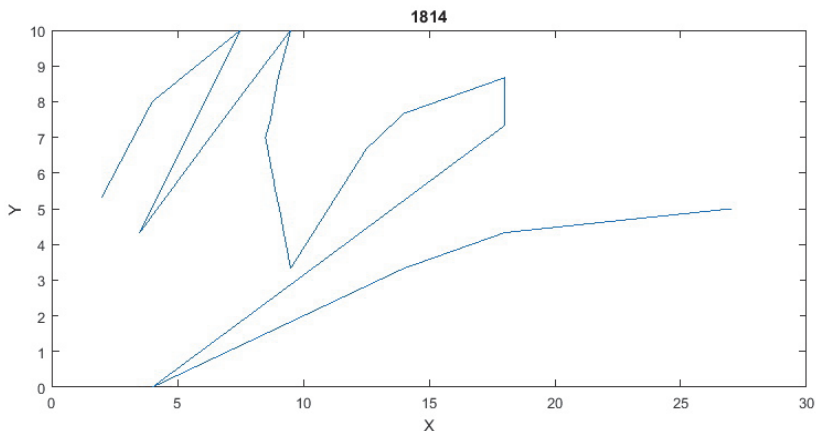


Рис. 5б. Изображение подписи Наполеона 1814 года

Расстояние между изображениями, обусловленное неравенствами границ баркодов нулевой размерности подписей 1809 и 1814 годов: $d_{W,2}(B, B') = 8,17$.

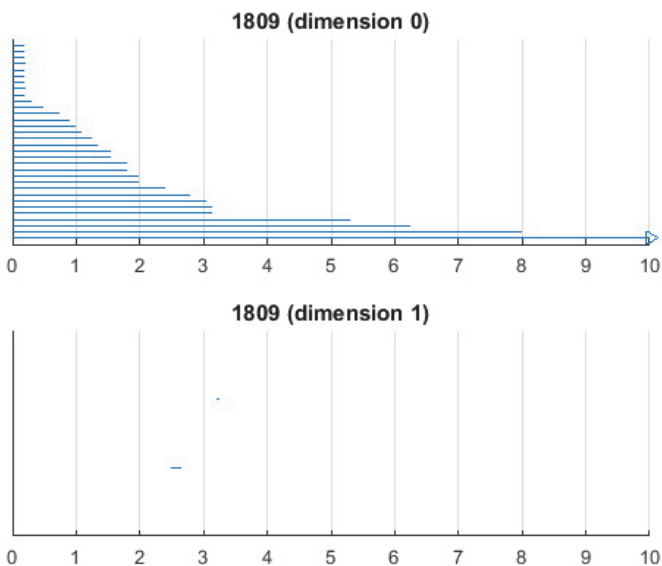


Рис. 6а. Баркоды персистентных гомологий изображения подписи Наполеона 1809 года

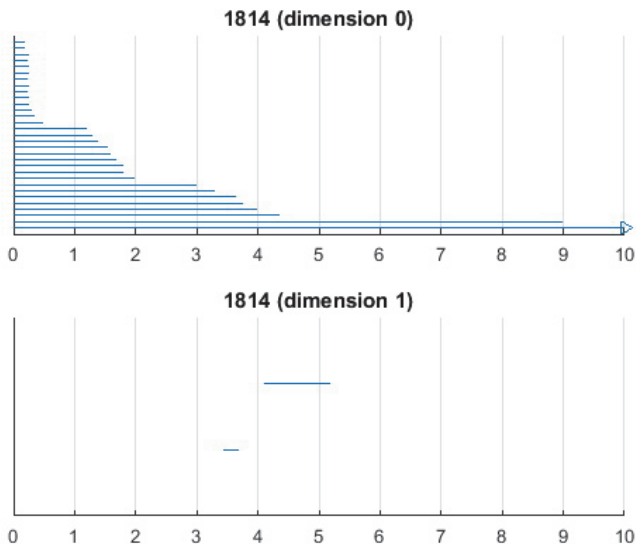


Рис. 6b. Баркоды персистентных гомологий изображения подписи Наполеона 1814 года

Поставим в соответствие каждому симплексу нулевой размерности (точке) значение функции — момент времени написания этой точки. Если обе подписи формировались равномерно (с нулевым ускорением) в течение 2 секунд, то расстояние между изображениями, обусловленное неравенствами функций исходного и терминального изображений: $d_f(B, B') = 0$. Модифицированное расстояние Вассерштейна между изображениями подписей 1809 и 1814 годов при $\mu_b = 1, \mu_f = 1$ в этом случае равно расстоянию между изображениями, обусловленному неравенствами границ баркодов: $d_{MW}(\mu_b, \mu_f, B, B') = d_{W,2}(B, B') = 8,17$. Если подпись 1809 года формировалась равномерно (с нулевым ускорением) в течение 2 секунд (скорость 0,0645 точек/с), а подпись 1814 года формировалась ускоренно (с ускорением 0,004162 точек/с²) в течение 2 секунд (начальная скорость 0,002 точек/с), то расстояние между изображениями, обусловленное неравенствами функций исходного и терминального изображений составляет: $d_f(B, B') = 2,03$. Модифицированное расстояние Вассерштейна между изображениями подписей 1809 и 1814 годов при $\mu_b = 1, \mu_f = 1$ в этом случае равно:

$$d_{MW}(\mu_b, \mu_f, B, B') = (\mu_b \cdot d_{W,2}(B, B') + \mu_f \cdot d_f(B, B'))^{1/2} = 8,42. \quad \square$$

Сравним метод на основе построения минимизируемого функционала и метод на основе нахождения расстояния Вассерштейна между баркодами.

Превосходство метода сравнения изображений на основе построения минимизируемого функционала заключается в следующем: возможность построения оптимальной траектории диффеоморфного отображения с формированием любого количества промежуточных форм; возможность построения метаморфозиса, при котором исходное и терминальное изображения не являются диффеоморфными и имеют разные топологические характеристики. Основным недостатком метода является низкая эффективность обучения: при количестве точек изображений более 100 длительность процесса обучения может превысить 1000 секунд (при использовании CPU Intel Pentium Dual-Core).

Преимущества метода сравнения изображений на основе нахождения расстояния Вассерштейна между баркодами заключаются в следующем: инвариантность метода по отношению к вращению и переносу изображений; повышение эффективности по отношению к методу на основе построения минимизируемого функционала на два порядка, если эффективность определяется количеством вычислительных операций или длительности решения задачи на ЭВМ. При количестве точек изображений, равном 16, и делении кривой изображения на 200 отрезков длительности решения задачи составляет 1-2 секунды. Основной недостаток метода — невозможность определения траектории эволюции изображения от начального к терминальному.

6. Формирование сглаженного коцикла. Гладкость функций от симплексов может быть измерена лапласианом. Рассмотрим граф, ассоциированный с X_k , где каждый k -симплекс представлен узлом и существует ребро, если два k -симплекса имеют непустое пересечение. Определим элементы весовой матрицы $W = (w_{ij})$ как:

$$w_{ij} = \begin{cases} v(\sigma_i)v(\sigma_j), & \text{если } \sigma_i \cap \sigma_j \neq \emptyset, \\ 0, & \text{иначе,} \end{cases}$$

где $v(\sigma_i)$ — размер σ_i , например площадь 2-симплекса. Взвешенный граф лапласиан графа \mathcal{L}_W определяется из соотношения [19]:

$$\mathcal{L}_{w_i, j} = \begin{cases} 1 - \frac{w_{ij}}{w_i}, & \text{если } i = j, w_i \neq 0, \\ -\frac{w_{ij}}{\sqrt{w_i w_j}}, & \text{если } \sigma_i \cap \sigma_j \neq \emptyset, i \neq j, \\ 0, & \text{иначе,} \end{cases} \quad (6)$$

где $w_i = \sum_j w_{ij}$. Гладкий коцикл $\bar{\omega}$ может быть получен решением задачи:

$$\bar{\alpha} = \arg \min_{\alpha \in C^{k-1}(X, \mathbb{R})} \|\mathcal{L}_W(\hat{\omega} + d\alpha)\|^2. \quad (7)$$

Коцикл $\hat{\omega} \in C^k(X, \mathbb{R})$ не всегда является точным. На основе построения гладкого точного коцикла $\bar{\omega} = \hat{\omega} + d\alpha \in C^k(X, \mathbb{R})$ симплициального комплекса решением задачи (7) можно построить такой коцикл $\bar{\alpha} \in C^{k-1}(X, \mathbb{R})$, что действие кограничного оператора d_{k-1} ($d_{k-1}: C^{k-1} \rightarrow C^k$) на $\bar{\alpha}$ приводит к $\bar{\omega}$: $d_{k-1}\bar{\alpha} = \bar{\omega}$.

Пример 4. Рассмотрим симплициальный комплекс X с пятью вершинами (v_0, v_1, v_2, v_3, v_4) и пятью ребрами (e_0, e_1, e_2, e_3, e_4) единичной длины. 1-коцепь $\omega_1 = [1, 0, 0, 0, 0]^T \in C^1(X, \mathbb{R})$ является вещественным коциклом; значения функций на ребрах: $\hat{\omega}(e_0) = 1$, $\hat{\omega}(e_1) = \hat{\omega}(e_2) = \hat{\omega}(e_3) = \hat{\omega}(e_4) = 0$. Так как $\omega_1 = [1, 0, 0, 0, 0]^T$ не является точной ($\omega \neq d_0\alpha; \forall \alpha \in C^0(X, \mathbb{R})$), то найдем такую 0-коцепь $\bar{\alpha} \in C^0(X, \mathbb{R})$, которая решает задачу минимизации (7). Весовая матрица равна:

$$w_{ij} = v(\sigma_i)v(\sigma_j) = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix}; w_0 = w_1 = w_2 = w_3 = w_4 = 3$$

при $v(\sigma_i) = 1, \forall i = 0, \dots, 4$. Матрица Лапласа \mathcal{L}_W , определенная из соотношения (6), равна:

$$\mathcal{L}_W = \frac{1}{3} \begin{bmatrix} 2 & -1 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix}.$$

Решая задачу (7), получим сглаженный коцикл $\bar{\omega}_1 = \hat{\omega}_1 + d\bar{\alpha}_1 = [0.2, 0.2, 0.2, 0.2, 0.2]^T$ с 0-коцепью $\bar{\alpha}_1 = [1.0, 0.2, 0.4, 0.6, 0.8]^T$; $d\bar{\alpha}_1 = [-0.8, 0.2, 0.2, 0.2, 0.2]^T$, который минимизирует $\|\mathcal{L}_W(\hat{\omega}_1 + d\bar{\alpha}_1)\|_2^2$, так как: $\mathcal{L}_W(\hat{\omega}_1 + d\bar{\alpha}_1) = [0, 0, 0, 0, 0]^T$, и $\|\mathcal{L}_W(\hat{\omega}_1 + d\bar{\alpha}_1)\|_2^2 = 0$.

Аналогично для 1-коцепи $\omega_2 = [0, 0, 0, 0, 1]^T \in C^1(X, \mathbb{R})$ получим 0-коцепь $\bar{\alpha}_2 = [0.2, 0.4, 0.6, 0.8, 1.0]^T$. Расстояние между ω_1 и ω_2 (см. (4)): $d_{f_1} = \sqrt{2}$; между $\bar{\alpha}_1$ и $\bar{\alpha}_2$: $d_{f_0} = \sqrt{0.8}$. \square

7. Заключение. Рассмотрена задача сравнения начального и терминального изображений, которая решается на основе построения минимизируемого функционала, характеризующего эволюцию диффеоморфного преобразования изображения от начального до терминального, и штрафа за отклонение траектории изображения от требуемой траектории.

В работе применяются алгоритмы построения множеств полуинтервалов вещественных баркодов персистентных гомотопий для сравнения диффеоморфных изображений объектов. Использование персистентных гомотопий позволяет повысить количество информации о форме объекта.

Достоинство метода сравнения изображений на основе построения минимизируемого функционала заключается в возможности построения оптимальной траектории диффеоморфного отображения. Преимущество метода сравнения изображений на основе нахождения расстояния Вассерштейна между баркодами определяется инвариантностью метода по отношению к вращению и переносу изображений и повышению эффективности по отношению к методу на

основе построения минимизируемого функционала на два-три порядка, если эффективность определяется количеством вычислительных операций.

Для внедрения как геометрической, так и негеометрической информации в единые топологические описания предлагается метод, основанный на персистентных когомологиях. Баркоды персистентных гомологий расширяются функциями от симплексов для представления разнородной информации.

Для нахождения расстояния между изображениями при распознавании образов предлагается модифицировать метод Вассерштейна и в формулу для расстояния между изображениями, определяемого по баркодам персистентных гомологий, ввести негеометрическую информацию о расстояния между изображениями.

Важным направлением применения персистентных (ко)гомогий является изучение инвариантов больших объемов данных.

Литература

1. *Trounev A., Younes L.* Metamorphoses through lie group action // *Foundations of Computation Mathematics.* 2005. vol. 5. no. 2. pp. 173–198.
2. *Charon N., Charlier B., Trounev A.* Metamorphoses of functional shapes in Sobolev spaces // 2016. arXiv preprint arXiv: 1608.01832. 46 p.
3. *Younes L., Arrate F., Miller M.I.* Evolutions equations in computational anatomy // *NeuroImage.* 2009. vol. 45. no. 1. pp. S40–S50.
4. *Younes L.* Diffeomorphic Learning // 2018. arXiv preprint arXiv: 1806.01240. 17 p.
5. *Zhang M., Fletcher P.T.* Fast Diffeomorphic Image Registration via Fourier-Approximated Lie Algebras // *International Journal of Computer Vision.* 2018. vol. 127. no. 1. pp. 61–73.
6. *Marsland S., McLachlan R.I.* A Hamiltonian particle method for diffeomorphic image registration // *Biennial International Conference on Information Processing in Medical Imaging.* 2007. pp. 396–407.
7. *Carlsson G.* Topological pattern recognition for point cloud data // *Acta Numerica.* 2014. vol. 23. pp. 289–368.
8. *Kerber M.* Persistent Homology: State of the Art and Challenges // *International Mathematische Nachrichten.* 2016. vol. 231. no. 15-33. pp. 1.
9. *Hatcher A.* Algebraic topology // Cambridge UP. 2002. 544 p.
10. *Deo S.* Algebraic topology // Springer Nature Singapore Private Limited. 2018. 351 p.
11. *Leichter S.V., Chukanov S.N.* Comparing images based on their diffeomorphic transformation // *Computer Optics.* 2018. vol. 42. no. 1. pp. 96–104.
12. *Durrleman S. et al.* Deformetrics: morphometry of shape complexes with space deformations // *Neuroimage.* 2014. vol. 101. pp. 35–49.
13. *Chukanov S.N., Ulyanov D.V.* Formation of invariants in the visualization of vector fields based on the construction of the homotopy operator // *Computer Optics.* 2012. vol. 36. no. 4. pp. 622–626.
14. *Chukanov S.N.* Constructing invariants for visualization of vector fields defined by integral curves of dynamic systems // *Optoelectronics, Instrumentation and Data Processing.* 2011. vol. 47. no. 2. pp. 151–155.
15. *Edelsbrunner H.* A Short Course in Computational Geometry and Topology // Springer. 2014. 110 p.

16. *Edelsbrunner H., Virk Z., Wagner H.* Topological Data Analysis in Information Space // 2019. arXiv preprint arXiv: 1903.08510.pdf. 18 p.
17. *Potierovich L., Rosen D., Samvelyan K., Zhang J.* Topological Persistence in Geometry and Analysis // 2016. arXiv preprint arXiv: 1904.04044.pdf. 127 p.
18. *Adams H., Tausz A.* JavaPlex Tutorial. URL: http://www.math.colostate.edu/~adams/research/javaplex_tutorial.pdf. 2011. 20 p. (дата обращения: 10.07.2018).
19. *Chung F.R., Graham F.C.* Spectral graph theory // American Mathematical Society. 1997. 207 p.
20. *Cang Z., Wei G.* Persistent cohomology for data with multicomponent heterogeneous information. 2018. arXiv preprint arXiv: 1807.11120. 25 p.
21. *de Silva V., Morozov D., Vejdemo-Johansson M.* Persistent cohomology and circular coordinates // Discrete and Computational Geometry. 2011. vol. 45. no. 4. pp. 737–759.

Чуканов Сергей Николаевич — д-р техн. наук, профессор, ведущий научный сотрудник, лаборатория методов преобразования и представления информации, Федеральное государственное бюджетное учреждение науки Институт математики им. С.Л. Соболева СО РАН, Омский филиал (ОФ ИМ СО РАН). Область научных интересов: системы управления летательными аппаратами, распознавание образов. Число научных публикаций — 78. ch_sn@mail.ru; ул. Певцова, 13, 644043, Омск, Российская Федерация; р.т.: +7(3812)236-739; факс: +7(3812) 23-45-84.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проекты № №18–07–00526 и 18–08–01284) и при поддержке программы фундаментальных научных исследований СО РАН № I.5.1., проект № 0314-2019-0020.

S.N. CHUKANOV
**COMPARISON OF OBJECTS' IMAGES BASED ON
COMPUTATIONAL TOPOLOGY METHODS**

Chukanov S.N. Comparison of Objects' Images based on Computational Topology Methods.

Abstract. The paper considers methods for comparison of objects' images represented by sets of points using computational topology methods. The algorithms for construction of sets of real barcodes for comparison of objects' images are proposed. The determination of barcodes of object forms allows us to study continuous and discrete structures, making it useful in computational topology. A distinctive feature of the use of the proposed comparison methods versus the methods of algebraic topology is obtaining more information about objects' form. An important area of application of real-valued barcodes is studying invariants of big data. Proposed method combines the technology of barcodes construction with embedded non-geometrical information (color, time of formation, pen pressure), represented as functions of simplicial complexes. To do this, barcodes are expanded with functions from simplexes to represent heterogeneous information. The proposed structure of extended barcodes increases the effectiveness of persistent homology methods when comparing images and pattern recognition. A modification of the Wasserstein method is proposed for finding the distance between images by introducing non-geometric information about the distances between images, due to inequalities of the functions of the source and terminal images of the corresponding simplexes. The geometric characteristics of an object can change with diffeomorphic deformations; the proposed algorithms for the formation of expanded image barcodes are invariant to rotation and translation transformations. We considered a method for determining the distance between sets of points representing the curves, taking into account an orientation of curves' segments. The article is intended for a reader who is familiar with basic concepts of algebraic and computational topology, the theory of Lie groups, and diffeomorphic transformations.

Keywords: Pattern Recognition, Diffeomorphic Transformations, Persistent Homology, Wasserstein Distance.

Chukanov Sergey Nikolayevich — Ph.D., Dr.Sci., Professor, Leading Researcher, Laboratory of Methods for Transformation and Representation of Information, Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences. Research interests: spacecraft control systems, pattern recognition. The number of publications — 78. ch_sn@mail.ru; 13, Pevtsova str., 644043, Omsk, Russian Federation; office phone: +7(3812)236-739; fax: +7(3812) 23-45-84.

Acknowledgements. The research is supported by the RFBR (projects no. 18–07–00526, no. 18–08–01284) and the Basic Research Program of the Siberian Branch of the Russian Academy of Sciences No. I.5.1., project No. 0314-2019-0020.

References

1. Trouvé A., Younes L. Metamorphoses through lie group action. *Foundations of Computation Mathematics*. 2005. vol. 5. no. 2. pp. 173–198.
2. Charon N., Charlier B., Trouvé A. Metamorphoses of functional shapes in Sobolev spaces. 2016. arXiv preprint arXiv: 1608.01832. 46 p.
3. Younes L., Arrate F., Miller M.I. Evolutions equations in computational anatomy. *NeuroImage*. 2009. vol. 45. no. 1. pp. S40–S50.
4. Younes L. Diffeomorphic Learning. 2018. arXiv preprint arXiv: 1806.01240. 17 p.

5. Zhang M., Fletcher P.T. Fast Diffeomorphic Image Registration via Fourier-Approximated Lie Algebras. *International Journal of Computer Vision*. 2018. vol. 127. no. 1. pp. 61–73.
6. Marsland S., McLachlan R.I. A Hamiltonian particle method for diffeomorphic image registration. Biennial International Conference on Information Processing in Medical Imaging. 2007. pp. 396–407.
7. Carlsson G. Topological pattern recognition for point cloud data. *Acta Numerica*. 2014. vol. 23. pp. 289–368.
8. Kerber M. Persistent Homology: State of the Art and Challenges. *International Mathematische Nachrichten*. 2016. vol. 231. no. 15-33. pp. 1.
9. Hatcher A. Algebraic topology. Cambridge UP. 2002. 544 p.
10. Deo S. Algebraic topology. Springer Nature Singapore Private Limited. 2018. 351 p.
11. Leichter S.V., Chukanov S.N. Comparing images based on their diffeomorphic transformation. *Computer Optics*. 2018. vol. 42. no. 1. pp. 96–104.
12. Durrleman S. et al. Deformetrics: morphometry of shape complexes with space deformations. *Neuroimage*. 2014. vol. 101. pp. 35–49.
13. Chukanov S.N., Ulyanov D.V. Formation of invariants in the visualization of vector fields based on the construction of the homotopy operator. *Computer Optics*. 2012. vol. 36. no. 4. pp. 622–626.
14. Chukanov S.N. Constructing invariants for visualization of vector fields defined by integral curves of dynamic systems. *Optoelectronics, Instrumentation and Data Processing*. 2011. vol. 47. no. 2. pp. 151–155.
15. Edelsbrunner H. A Short Course in Computational Geometry and Topology. Springer. 2014. 110 p.
16. Edelsbrunner H., Virk Z., Wagner H. Topological Data Analysis in Information Space. 2019. arXiv preprint arXiv: 1903.08510.pdf. 18 p.
17. Polterovich L., Rosen D., Samvelyan K., Zhang J. Topological Persistence in Geometry and Analysis. 2016. arXiv preprint arXiv: 1904.04044.pdf. 127 p.
18. Adams H., Tausz A. JavaPlex Tutorial. URL: http://www.math.colostate.edu/~adams/research/javaplex_tutorial.pdf. 2011. 20 p. (дата обращения: 10.07.2018).
19. Chung F.R., Graham F.C. Spectral graph theory. American Mathematical Society. 1997. 207 p.
20. Cang Z., Wei G. Persistent cohomology for data with multicomponent heterogeneous information. 2018. arXiv preprint arXiv: 1807.11120. 25 p.
21. de Silva V., Morozov D., Vejdemo-Johansson M. Persistent cohomology and circular coordinates. *Discrete and Computational Geometry*. 2011. vol. 45. no. 4. pp. 737–759.

**Ю.И. СЕНКЕВИЧ, Ю.В. МАРАПУЛЕЦ, О.О. ЛУКОВЕНКОВА, А.А. СОЛОДЧУК
МЕТОДИКА ВЫДЕЛЕНИЯ ИНФОРМАТИВНЫХ ПРИЗНАКОВ В
СИГНАЛАХ ГЕОАКУСТИЧЕСКОЙ ЭМИССИИ**

Сенкевич Ю.И., Марапудец Ю.В., Луковенкова О.О., Солодчук А.А. Методика выделения информативных признаков в сигналах геоакустической эмиссии.

Аннотация. Исследования геоакустической эмиссии в сейсмоактивном регионе на Камчатке показывают, что при подготовке землетрясений и последующей релаксации поля локальных напряжений в пункте наблюдений в геоакустических сигналах возникают ярко выраженные импульсные аномалии. Качественному выделению таких аномалий препятствуют сильное искажение и ослабление амплитуды сигнала. Обзор существующих методов анализа акустической эмиссии показывает, что чаще всего исследователи обращаются к анализу энергетических и статистических свойств сигналов, как более доступных для изучения. Отличительными особенностями предлагаемого авторами подхода являются выделение информативных признаков на основе анализа временной и частотно-временной структур геоакустических сигналов и описание многообразных форм распознаваемых импульсов ограниченным набором паттернов. Настоящее исследование открывает перспективу разработки новых методов выявления аномального поведения геоакустических сигналов, в том числе и перед землетрясениями.

В работе описана методика извлечения информации из потоков импульсов геоакустической эмиссии звукового частотного диапазона. Представлена математическая модель геоакустического импульса, отражающая процесс генерации сигнала от множества элементарных источников. Приводится решение задачи выделения информативных признаков в геоакустических сигналах путем описания фрагментов сигнала матрицами отношений амплитуд локальных экстремумов и интервалов между ними. Приводится результат применения разработанного алгоритма для автоматического описания структуры выделяемых импульсов и для образования множества паттернов, характеризующих особенности сигналов геоакустической эмиссии, наблюдаемых на полевых станциях ИКИР ДВО РАН. Представлена методика сокращения размерности множества выделенных импульсов, позволяющая найти близкие по структуре паттерны. Предложено решение проблемы обработки большого потока данных путем унификации описания импульсов и их систематизации. Результаты проведенных научных работ позволили создать инструмент для исследования динамических свойств сигналов геоакустической эмиссии в интересах разработки детекторов предсказания землетрясений.

Ключевые слова: геоакустическая эмиссия, модель геоакустического импульса, разреженная аппроксимация, информативные признаки, паттерны сигнала.

1. Введение. Геофизические процессы, связанные с подвижками горных пород и грунтов, сопровождаются акустическими колебаниями. Такие сигналы отличаются многообразием форм, широким динамическим диапазоном частот и амплитуд. Многолетние исследования [1-4] позволили выявить закономерные изменения характеристик этих сигналов, которые связаны с процессами развития и подготовки землетрясений.

Наблюдения на Камчатке, проводящиеся более 15 лет, показывают, что в течение нескольких суток перед сильными землетрясениями на

расстоянии сотен километров от эпицентра возникают аномалии сигналов геоакустической эмиссии в звуковом диапазоне частот (ГАЭ), выражающиеся в виде резкого и продолжительного повышения интенсивности акустических сигналов. Впервые наличие таких аномалий в диапазоне частот от единиц до десятков килогерц обсуждалось в работе А.В. Купцова [5]. В дальнейших исследованиях [6-9] эти результаты нашли подтверждение и были дополнены. Непрерывная регистрация геоакустической эмиссии на полигонах ИКИР ДВО РАН приводит к стремительному накоплению огромных объемов необработанных данных. В связи с чем возникает проблема обработки большого потока данных, требующая применения специальных методов анализа, хранения и передачи информации. Одним из способов решения данной проблемы является унификация описания и систематизация импульсов с целью выявления характерных и аномальных паттернов и их дальнейшего анализа в привязке к ряду сейсмических событий.

2. Постановка задачи исследования. Сигналы ГАЭ представляют собой упругие колебания, возникающие в процессе локальной динамической перестройки внутренней структуры приповерхностных осадочных пород, которые характеризуются малой прочностью и высокой пластичностью. В них акустическая эмиссия возникает даже при слабых деформационных изменениях, соизмеримых по уровню с приливными.

Распространение упругих волн в осадочных породах имеет характерные особенности и существенно отличается от распространения в однородных изотропных твердых средах. Модель среды для осадочных пород из-за ее сложности в настоящее время полностью не сформирована. Наиболее близкой является модель пористо-упругой, насыщенной жидкостью или газом среды. Ее основы были заложены в работах М.А. Био [10] и Я.И. Френкеля [11]. Теория Био—Френкеля базируется на предположении, что пористо-упругая среда состоит из упругого однородного скелета с сообщающимися между собой порами, которые заполнены вязкой сжимаемой жидкостью. В отличие от однородных изотропных сред возмущенное состояние такой среды характеризуется двумя полями осредненных векторов смещений: для жесткого пористого каркаса и для жидкости в поровом пространстве относительно каркаса. При анализе распространения упругих волн в осадочных породах с частотами до 20-30 кГц можно воспользоваться более простой моделью пористо-упругой среды Ф. Гассмана, согласно которой каркас и флюид движутся вместе, без проскальзывания [12]. Кроме того, в случае размещения акустического приемника в водоеме у дна [5] можно ограничиваться исследованием только продольных волн, так как поперечные волны в воде

не распространяются, а поверхностные волны на границе между жидкостью и донными породами сильно затухают.

Измерения в осадочных породах, проводимые на полевых станциях ИКИР ДВО РАН, показывают, что типичный сигнал ГАЭ представляет собой последовательность импульсов различной амплитуды и длительности с ударным возбуждением и частотным заполнением от сотен герц до первых десятков килогерц [13]. Форма импульсов геоакустической эмиссии зависит от природы процесса и свойств среды. Так, при уплотнении и разуплотнении осадочных пород, сопровождающихся трением и взаимным соударением входящих в них неоднородных фракций, характерны импульсы с быстро нарастающим фронтом и затяжным спадом. Разнообразие акустических импульсов продиктовано наложением в точке приема сигналов от множества источников, возникающих в объеме пространства, определяемом чувствительностью приемной аппаратуры, а также влиянием возможных шумов техногенного и естественного происхождения.

В настоящее время ведется поиск информации, содержащейся в многообразных динамически изменяющихся структурных формах сигналов ГАЭ. Разработке методов анализа сигналов ГАЭ посвящено достаточно большое количество исследований [13-16]. Достижению значимых результатов в этой области знаний мешают некоторые ограничения. Так, качественному выделению импульсов в сигналах ГАЭ препятствуют их сильное искажение и ослабление амплитуды, связанные с интерференцией сигналов от множества распределенных в зоне приема источников, а также с наличием в ней разнородных отличающихся по плотности фракций. Кроме того, при удалении от точки приема в среде начинают сказываться эффекты неравномерности фазовых скоростей в импульсах на разных частотах, искажающие их форму. Анализ известных методов показывает, что в большинстве случаев исследователи обращаются к анализу энергетических и статистических свойств сигналов ГАЭ, как более доступных для изучения. Отличительной особенностью настоящего исследования является разработка эффективного инструмента, позволяющего изучать многообразие форм распознаваемых импульсов ГАЭ. Целесообразность же проводимых исследований связана с перспективой разработки новых методов выявления аномального поведения геоакустических сигналов, в том числе и перед землетрясениями.

3. Представление модели. Учитывая, что геоакустические импульсы представляют собой аддитивные сигналы, которые образуются наложением элементарных моночастотных импульсов друг на друга, предлагается следующая модель геоакустического сигнала:

$$\begin{cases} s(t) = \sum_{m=1}^{N_1} \alpha_m g_m(t) + \sum_{n=1}^{N_2} \beta_n g_n(t) + R(t), \\ \|R(t)\| \rightarrow \min, \end{cases} \quad (1)$$

где $s(t)$ — фрагмент сигнала, включающий импульс; $g_m(t)$ — функции, аппроксимирующие импульс; $g_n(t)$ — функции, аппроксимирующие шумовую составляющую импульса; α_m и β_n — коэффициенты аппроксимации; N_1 — количество функций, на которые раскладывается импульс; N_2 — количество функций, на которые раскладывается шумовая составляющая фрагмента сигнала. Невязка $R(t)$ определяет отклонение аппроксимации от исходного сигнала. Функциональные зависимости (1) с достаточной точностью описывают сигналы от элементарных источников, а сама модель включает ограниченное число компонент, содержащих информацию о количественных и качественных изменениях состояний среды.

4. Методика исследования. Согласно принятой модели (1) для обнаружения изменений состояний среды необходимо идентифицировать информативные признаки сигналов ГАЭ. Для этого была разработана методика, состоящая из четырех этапов:

- выделение участков сигнала, содержащих полезные данные (импульсы);
- анализ амплитудно-фазовой структуры сигналов;
- выделение базового множества импульсов ГАЭ;
- уточнение модели геоакустического импульса.

4.1. Выделение информативных участков сигнала. Для выделения одиночных импульсов в сигналах геоакустической эмиссии предлагается алгоритм, основанный на вычислении адаптивного порога. Расчет пороговых значений ведется в непересекающихся окнах фиксированной длины n на основе среднеквадратического отклонения (СКО) по формуле:

$$S_k = S_{k-1} + A \cdot (\sigma_k - \sigma_{k-1}),$$

где S_k — пороговое значение в окне от $(k \cdot n)$ -го до $(k \cdot n + n - 1)$ -го отсчета, σ_i — СКО, вычисленное в окне от $((i - 1) \cdot n)$ -го до $(i \cdot n - 1)$ -го отсчета, A — экспериментально определяемый параметр. Поскольку вычисляемый порог должен зависеть лишь от уровня фонового шума, участки сигнала, содержащие импульсы, были исключены из процесса вычисления пороговой функции.

Для определения границ импульса сигнал просматривается скользящим окном длиной Δ в прямом и обратном направлениях, начиная с точки превышения порогового значения. Регистрация начала и конца

импульса производится в момент, когда все отсчеты в окне Δ окажутся меньше заданного значения S_0 :

$$S_0 = B \cdot \sigma_{\Delta},$$

где σ_{Δ} — СКО на данном участке сигнала, B — экспериментально определяемый параметр. Длина окна Δ также определяется экспериментально.

Выбор значения параметра B порогового детектора в условиях стационарного характера и нормального распределения шума осуществляется по известному в статистике правилу «трех сигм». Согласно данному правилу вероятность попадания значений шума в интервал $\mu \pm 3\sigma$ составляет 0.9973. Однако в условиях нестационарных помех, с которыми приходится иметь дело на практике, выбор фиксированного значения эмпирического параметра B становится нетривиальной задачей. Вычислительные эксперименты показали, что при заниженных значениях параметра B ($0 \leq B \leq 1.5$) и сильных отклонениях (выбросах) значений шума резко возрастает вероятность ошибок второго рода — выявление ложных импульсов, а при завышенных значениях B ($B > 1.5$) увеличивается вероятность ошибок первого рода, начинаются пропуски импульсов с небольшими значениями амплитуды.

Изучение специфических свойств геоакустических сигналов позволило снизить пороговые значения детектора импульсов на 3-5 дБ при фиксированной вероятности корректного выделения импульсов. Разработанный алгоритм детектора назван логической фильтрацией. Его главная идея состоит в том, что выделенный пороговой схемой детектора фрагмент сигнала проверяется на целостность цуга локальных максимумов и минимумов, зафиксированных на его протяжении. Если длина непрерывной последовательности обнаруженных в сигнале локальных экстремумов будет превышать заданную величину N , то данный фрагмент сигнала принимается за искомый импульс. И наоборот, если число непрерывно следующих друг за другом локальных экстремумов меньше установленного порога N , отсчеты фрагмента сигнала приравняются нулю. Этот несложный алгоритм оказывается весьма эффективным в силу физических особенностей генерации импульсов ГАЭ, поскольку неоднородные по плотности фракции осадочных пород под внешним силовым воздействием, как уже отмечалось ранее, становятся источниками колебаний различной частоты, продолжительности и амплитуды, а специфика предложенного алгоритма не влияет на подсчет числа экстремумов в выделяемых фрагментах сигнала. Таким образом, алгоритм инвариантен трансформациям

сигнала в порождающей неоднородной среде. Более детально алгоритм логического фильтра представлен в виде блок-схемы на рисунке 1.

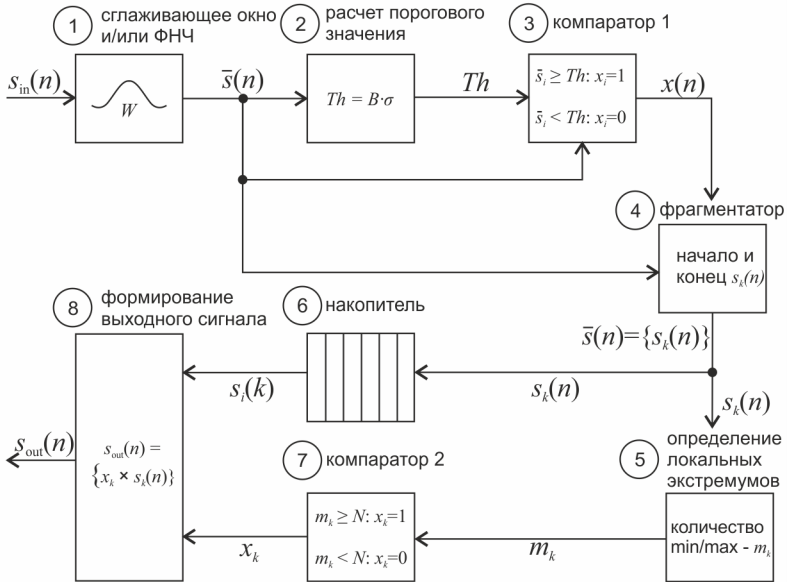


Рис. 1. Блок-схема алгоритма логического фильтра

Алгоритм работы фильтра следующий:

– оцифрованный сигнал ГАЭ $s_{in}(n)$ с частотой дискретизации 48000 Гц обрабатывается сглаживающим временным окном протяженностью от 8 до 16 отсчетов и/или низкочастотным фильтром с граничной частотой, выбираемой в зависимости от задачи исследования (от 500 Гц до 16000 Гц). Данная операция позволяет подавить высокочастотную составляющую шума и уменьшить влияние его одиночных выбросов (блок 1);

– для отфильтрованного сигнала $\bar{s}(n)$ непрерывно рассчитывается порог Th на основе СКО сигнала (блок 2) во временном окне длиной, на порядок превышающей протяженность самых длительных ожидаемых импульсов (как правило, от 3000 до 5000 отсчетов оцифрованного сигнала);

– пороговое значение Th сравнивается со значениями отсчетов отфильтрованного сигнала, затем на основе результатов сравнений формируется управляющий сигнал $x(n)$ (блок 3), при этом $x_i = 1$, если значения

соответствующих отсчетов сигнала $\bar{s}(n)$ превышают пороговое значение B , и $x_i = 0$ в противном случае;

– сигналы $x(n)$ и $\bar{s}(n)$ с выхода линии задержки поступают в блок 4, выполняющий функцию вырезания фрагментов сигнала ГАЭ в соответствии со значениями управляющего сигнала: если $x_i \neq x_{i+1}$, то, начиная с \bar{s}_{i+1} , формируется новый фрагмент;

– полученные фрагменты сигнала ГАЭ $\{s_k(n)\}$ поочередно поступают на вход блока 5, где осуществляется определение локальных экстремумов (максимумов и минимумов) каждого фрагмента $s_k(n)$ с подсчетом их количества m_k ;

– количество локальных экстремумов m_k сравнивается с установленным пороговым значением N длины цуга локальных экстремумов (блок 7), таким образом принимается решение об идентификации фрагмента как импульса ГАЭ или о ложном срабатывании детектора;

– в первом случае фрагмент $s_k(n)$ с накопителя, где фрагменты сигнала сохраняются до момента принятия решения (блок 6), остается без изменения, во втором случае формируется нулевой сигнал, совпадающий по длительности с пропускаемым фрагментом, чем достигается непрерывность последовательности отсчетов выходного сигнала $s_{out}(k)$ (блок 8).

На рисунке 2 проиллюстрированы ключевые моменты логической фильтрации тестового сигнала, представляющего собой последовательность импульсов Берлаге с наложенным белым шумом. Параметры компьютерного моделирования:

- фильтрация сигнала не применялась;
- длительность генерируемых импульсов Берлаге — 208 отсчетов;
- максимальная амплитуда импульсов Берлаге — 100;
- относительная частота модуляции импульсов Берлаге — 0.18;
- коэффициент нарастания импульса Берлаге — 2.3;
- коэффициент затухания импульсов Берлаге — 0.16;
- заданная скважность импульсов — 5.8;
- СКО шума — 100;
- ширина окна расчета СКО — 2048 отсчетов;
- коэффициент порогового значения — $1.2 (Th = 1.2\sigma)$;
- расчетное отношение сигнал/шум (SNR) — 0.47 дБ;
- длина цуга локальных экстремумов $N = 10$.

Для тестирования алгоритма логической фильтрации, представленного на блок-схеме (рисунок 2), был проведен вычислительный эксперимент, результаты которого проиллюстрированы на рисунке 3. По вертикальной оси в логарифмическом масштабе отложено число коррект-

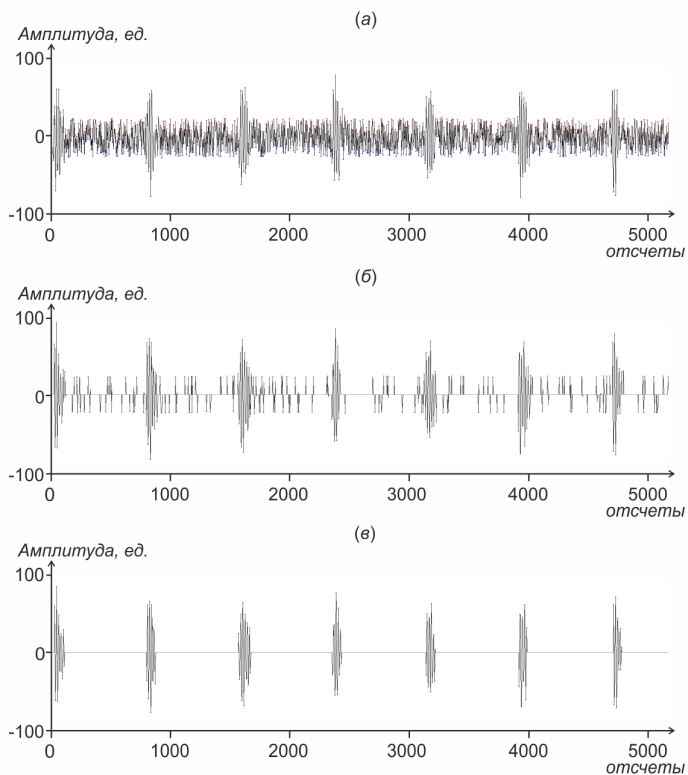


Рис. 2. Логическая фильтрация тестового сигнала: исходный сигнал (а); фрагментированный сигнал, выход блока 5 (б); результат обработки (в)

но выделенных импульсов ГАЭ, а по горизонтальной оси — значения SNR в дБ, рассчитанные во временном окне длиной 208 отсчетов.

Эксперимент проводился с помощью оригинальной компьютерной программы, имитирующей работу логического фильтра. Для эксперимента было сформировано 50 сигналов, содержащих по 100 импульсов Берлаге с характеристиками, описанными выше. Для получения сигналов с фиксированными значениями SNR от -0.5 до 20 дБ к импульсным сигналам добавлялся белый шум различной амплитуды. Далее подсчитывалось количество выделенных импульсов при применении различных значений коэффициента B . Порог Th менялся ступенчато от 0.5σ до 5.0σ , где σ — СКО накладываемого на сигнал белого шума.

На рисунке 3 представлены два типа зависимостей. Графики первого типа отражают результат применения простого порогового детектора (далее Д1) без логического фильтра к сигналу с неизвестными параметрами (на рисунке 3 обозначены сплошными линиями). Графики второго типа — результат включения в схему детектора логического фильтра (далее Д2, графики обозначены пунктирными линиями).

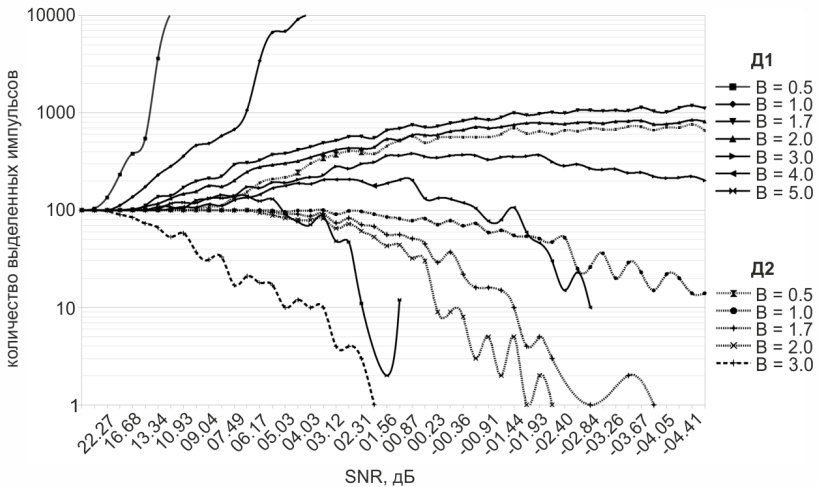


Рис. 3. Результаты сравнения пороговых детекторов без логических фильтров (Д1) и с логическими фильтрами (Д2)

Анализ графиков первой группы показывает, что детектор без логического фильтра (Д1) обеспечивает максимальную вероятность правильного обнаружения 0.95 при $B \geq 1.7$, что соответствует правилу «трех сигм». Дальнейшее уменьшение значения параметра B приводит к лавинному увеличению числа пропусков импульсов и к появлению значительного числа ложных срабатываний — ошибок первого и второго рода соответственно. Количество ошибок для обоих детекторов при различных значениях B представлено в таблице 1.

Детектор с логическим фильтром (Д2) тестировался при двух значениях $N = 12$ и $N = 3$ (минимальное количество локальных экстремумов в цуге). Данный подход позволил оценить чувствительность алгоритма к варьированию его внутренних параметров B и N . При повышении значения B все меньшее число локальных экстремумов импульса превышает порог Th , и следовательно, их число не превышает заданную величину N , тогда вероятность пропуска импульса увеличивается. Когда значение B

уменьшается длина последовательности локальных экстремумов шума может превысить параметр N логического фильтра, тогда вероятность ложной идентификации возрастает и возникают ошибки второго рода. В тех случаях, когда в ходе эксперимента размерность множества ложно выделенных импульсов превышала в три раза фактическое количество импульсов (100), возникла ситуация неопределенности в вычислении ошибки измерений, что и отмечено в таблице 1 (символом « \leftrightarrow »).

Таблица 1. Ошибки идентификации импульсов в вычислительном эксперименте для четырех значений SNR

Параметр B	Ошибки I рода		Ошибки II рода		SNR, дБ	σ_s^2/σ_n^2 , дБ
	D1	D2 ($N = 12/3$)	D1	D2 ($N = 12/3$)		
0.5	–	0/0	15283	0/–	10	5.3
1.0	–	0/0	363	0/63		
1.7	0	0/0	198	0/18		
2.0	0	0/0	156	0/0		
3.0	0	47/0	126	0/0		
0.5	–	0/0	15395	91/–	6.2	1.5
1.0	–	0/0	3325	0/224		
1.7	–	1/0	326	0/26		
2.0	0	6/0	277	0/14		
3.0	0	82/0	169	0/0		
0.5	–	0/0	14212	273/–	3.1	-1.5
1.0	–	9/0	12936	0/–		
1.7	–	26/0	519	0/74		
2.0	–	35/0	415	0/52		
3.0	3	96/3	281	0/0		
0.5	–	5/5	13031	461/461	0	-1.4
1.0	–	22/22	15871	2/2		
1.7	–	63/63	782	0/0		
2.0	–	81/81	645	0/0		
3.0	–	100/100	363	0/0		

Применение логического фильтра позволяет существенно снизить количество ошибок второго рода для значений параметра $B \leq 1.7$ при отношениях SNR от 3 дБ до 0 дБ. Цена такого результата — повышение вероятности ошибки первого рода. Однако на практике исключение ложных срабатываний позволяет отбирать для дальнейшего анализа именно участки полезного сигнала, что облегчает формирование структурных шаблонов импульсов ГАЭ.

Результаты вычислительного эксперимента доказывают применимость алгоритма логической фильтрации в задачах обнаружения импульсов и очистки импульсных сигналов от шума. Следует отметить, что на практике протяженность выделяемого цуга локальных экстремумов

колеблется от 8 до 200 и более. Этот показатель может служить мерой надежности при идентификации импульсов ГАЭ, поскольку появление фрагментов шума с протяженностью цуга локальных экстремумов более N становится маловероятно.

Перед дальнейшим анализом проводится предварительная обработка данных, включающая процедуру нормирования и сглаживания данных, которая позволяет уменьшить влияние высокочастотных помех на амплитудно-фазовые характеристики сигнала.

4.2. Анализ амплитудно-фазовой структуры сигналов. Для описания амплитудно-фазовой структуры выделенных импульсов ГАЭ было решено применить аппарат структурно-лингвистической обработки и анализа сигналов [17], адаптированный к особенностям геоакустических сигналов. Амплитудно-фазовую структуру выделенных импульсов ГАЭ было решено описать на основе отношений локальных экстремумов импульса. Данный подход позволяет представить сигнал в рамках замкнутого множества без использования пороговых значений уровней квантования амплитуды и периода дискретизации, возникающих в процессе оцифровки аналогового сигнала. Повышение точности измерений требует увеличения указанных характеристик оцифровки и порождает комбинаторно зависимое разнообразие допустимых значений сигнала. Избежать такой ситуации возможно, если трансформировать сигнал в последовательность его локальных экстремумов и интервалов между ними. Такое преобразование не кажется критичным, если абсолютные значения сигнала между его соседними локальными экстремумами можно не учитывать, что вполне оправдано на практике измерений большинства натуральных сигналов. Сущность преобразования продемонстрирована на рисунке 4.

Имеем исходное множество последовательных значений локальных экстремумов $\{x_i\}$, где $i = 1 \dots N$. Дополним это множество еще одним множеством с последовательными значениями интервалов времени τ_i , где $i = 1 \dots N - 1$, между локальными экстремумами. Вычислим отношения для каждого из экстремумов следующим образом:

$$r_{i,j} = \begin{cases} 1, & x_i > x_j \\ 0, & x_i \leq x_j \end{cases} \quad \omega_{i,j} = \begin{cases} 1, & \tau_i > \tau_j \\ 0, & \tau_i \leq \tau_j \end{cases},$$

где $r_{i,j}$ — результат логического сравнения i -го и j -го значений амплитуд экстремумов; $\omega_{i,j}$ — результат логического сравнения i -го и j -го значений интервалов между экстремумами. Упорядочим ряды таких отношений в виде квадратных матриц для отношений $r_{i,j}$ и $\omega_{i,j}$.

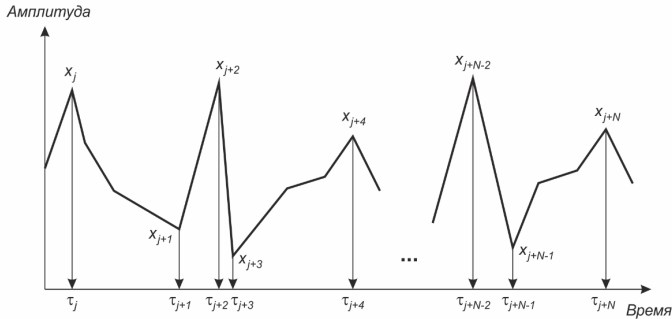


Рис. 4. Пример описания оцифрованного сигнала значениями локальных экстремумов и интервалов между ними

$$\mathbf{R} = \begin{pmatrix} 0 & r_{1,2} & \dots & r_{1,N-1} & r_{1,N} \\ r_{2,1} & 0 & & & r_{2,N} \\ \vdots & & \ddots & & \vdots \\ r_{N-1,1} & & & 0 & r_{N-1,N} \\ r_{N,1} & r_{N,2} & \dots & r_{N,N-1} & 0 \end{pmatrix}, \quad (2)$$

$$\mathbf{W} = \begin{pmatrix} 0 & \omega_{1,2} & \dots & \omega_{1,N-2} & \omega_{1,N-1} \\ \omega_{2,1} & 0 & & & \omega_{2,N-1} \\ \vdots & & \ddots & & \vdots \\ \omega_{N-2,1} & & & 0 & \omega_{N-2,N-1} \\ \omega_{N-1,1} & \omega_{N-1,2} & \dots & \omega_{N-1,N-2} & 0 \end{pmatrix}. \quad (3)$$

Полученные матрицы (2) и (3) являются избыточными в силу алгебраического свойства асимметрии неравенств (если $a > b$, то невозможно $b > a$), поэтому в дальнейшем будем использовать только верхние треугольные части каждой из построенных матриц, объединение которых образует одну квадратную матрицу порядка $N - 1$

$$\mathbf{RW} = \begin{pmatrix} r_{1,2} & r_{1,3} & \dots & r_{1,N-1} & r_{1,N} \\ \omega_{1,2} & r_{2,3} & & & r_{2,N} \\ \omega_{1,3} & \omega_{2,3} & \ddots & & \vdots \\ \vdots & & \ddots & r_{N-2,N-1} & r_{N-2,N} \\ \omega_{1,N-1} & \omega_{2,N-1} & \dots & \omega_{N-2,N-1} & r_{N-1,N} \end{pmatrix}. \quad (4)$$

Матрица (4) представляет описание участка сигнала (импульса), характеризующее взаимное положение локальных экстремумов. Преобразование сигнала, содержащего N экстремумов, в матрицу (4) обладает свойством инвариантности к операциям сдвига, а также амплитудного и временного транспонирования исходного сигнала. Инвариантность матрицы к изменениям сигнала вытекает из свойства неравенств: если $a > b$ и c – любое, то $a + c > b + c$ — для операции сдвига сигнала; если $a > b$ и $c > 0$, то $a \cdot c > b \cdot c$ — для операции транспонирования сигнала. В результате каждой матрице (4) можно сопоставить графический инвариант формы сигнала (паттерн).

Используя представленное преобразование, можно выполнить описание любого фрагмента сигнала. Пример такого структурного описания представлен на рисунке 5. Вверху слева показан эпизод оцифрованного сигнала с выделенным фрагментом; справа — матрица отношений выделенного фрагмента в символах «больше», «меньше», «равно»; внизу слева — паттерн выделенного фрагмента, построенный на основании матрицы отношений.

4.3. Методика выделения базового множества импульсов геоакустической эмиссии. Анализ импульсных сигналов ГАЭ проводится с целью получения информации о процессах, происходящих в литосфере. Классификация геоакустических сигналов, по существу, решает задачу классификации источников этих сигналов. Критерием классификации в предлагаемой авторами методике является структурное описание на основе локальных отношений экстремумов сигнала, которое как было показано ранее, может быть представлено матрицей амплитудно-фазовых отношений (4).

Согласно пункту 4.1 для выделения фрагментов сигнала, содержащих полезную информацию, используются пороговые схемы. Это приводит к накоплению огромного числа импульсов, описываемых разнообразными паттернами, которые образуют алфавит рассматриваемого фрагмента сигнала. Для выделения базового множества наиболее ти-

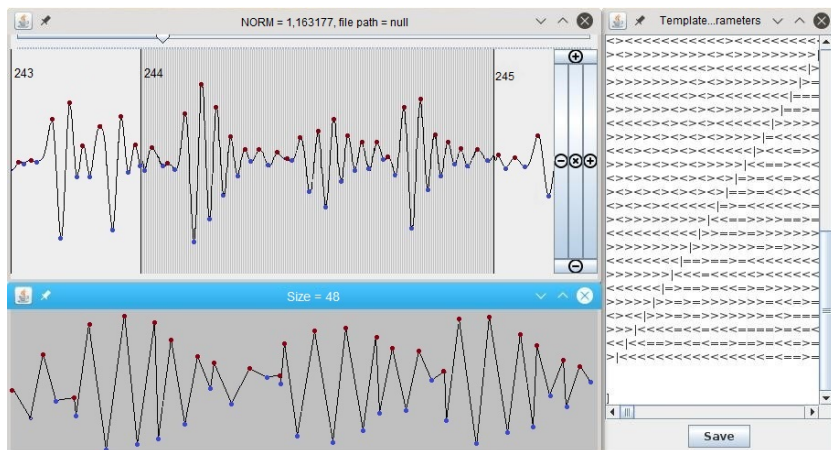


Рис. 5. Пример структурного описания фрагмента сигнала, полученного с помощью авторского программного обеспечения Registry

пичных форм импульсов ГАЭ необходимо сокращать объем алфавита. О причинах получаемого разнообразия импульсных сигналов сказано выше.

В основе механизма выделения базового множества импульсов ГАЭ лежит операция определения степени совпадения матриц структурного описания, соизмеримых по значениям порядка. Матрицы $\mathbf{A}_{n \times n}$ и $\mathbf{B}_{m \times m}$ ($n \geq m$), будут сравнимы, если $0 \leq (n - m)/n \leq P$, где P — эмпирически подбираемый порог.

Сравнение матриц одинакового порядка осуществляется путем подсчета совпадений их элементов при наложении матриц друг на друга. Если значения порядков отличны ($n > m$), осуществляется $n - m + 1$ сравнений, при этом на первом сравнении матрица порядка m вкладывается в матрицу порядка n таким образом, чтобы первые элементы первых строк совпадали, а для дальнейших сравнений сдвиг осуществляется вдоль диагонали большей матрицы, то есть матрица меньшего порядка смещается на один элемент вправо и вниз. По итогам $n - m + 1$ сравнения выбирается наибольшее количество совпадений. Данное значение сравнивается с пороговым, рассчитанным на основании эмпирически составленной формулы $S = g \cdot m^2$.

При превышении наибольшим количеством совпадений порога S принимается решение о допустимой близости структур сравниваемых импульсов. При этом статистика частоты появления импульса со структур-

ной матрицей большего порядка увеличивается на значение статистики частоты появления матрицы меньшего порядка, а паттерн с матрицей меньшего порядка удаляется из алфавита проанализированного фрагмента сигнала. Поясняющий пример сравнения двух матриц порядков 5 и 3 показан на рисунке 6.

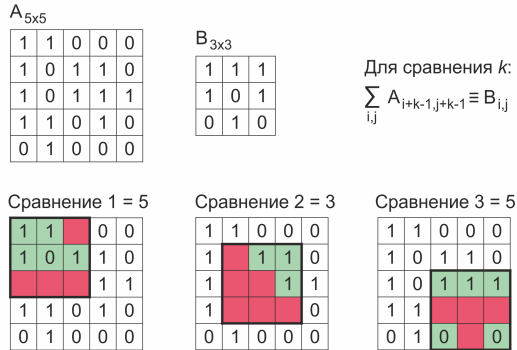


Рис. 6. Представление последовательности операций при выполнении алгоритма сравнения матриц и порядков 5 и 3 соответственно (пояснения в тексте)

4.4. Уточнение модели геоакустического импульса. На последнем этапе для полученного базового множества типичных геоакустических импульсов оцениваются параметры предлагаемой модели (1). В основу метода уточнения модели были положены идеи разреженной аппроксимации. Во-первых, разреженное представление строится по избыточной системе функций, принадлежащих разным семействам, что существенно повышает адаптивность модели импульса. Во-вторых, разреженные представления компактны и неизбыточны. Наконец, методы разреженной аппроксимации хорошо показали себя при решении аналогичных задач в смежных областях науки [18-20].

Решение задачи разреженной аппроксимации предполагает представление сигнала $s(t)$ в виде линейной комбинации минимально возможного числа функций $g_m(t)$ из предварительно выбранного базиса, что согласуется с требованиями к модели:

$$s(t) = \sum_{m=1}^N c_m g_m(t), \|c\|_0 \rightarrow \min, \tag{5}$$

где c — вектор коэффициентов разложения.

Точность и разреженность получаемых аппроксимаций напрямую зависит от выбранной системы функций (в терминах разреженной аппроксимации систему функций называют словарем, а каждую функцию в отдельности — частотно-временным атомом). Для анализа сигналов ГАЭ подходит словарь, построенный путем вариации параметров модулированных и сдвинутых функций Гаусса и Берлаге. Импульс Гаусса задается аналитическим выражением:

$$g(t) = A \cdot \exp(-B(t_{end}) \cdot \Delta \cdot t^2) \cdot \sin(2\pi ft),$$

где A — амплитуда, выбирается таким образом, чтобы $\|g(t)\| = 1$; t_{end} — длина атома; f — частота от 200 до 20000 Гц (диапазон регистрируемых частот); $B(t_{end})$ — предельное значение параметра B , вычисляемое по формуле:

$$B(t_{end}) = -\frac{4 \cdot \ln 0.05}{t_{end}^2},$$

Δ — коэффициент варьирования параметра B относительно предельного значения.

Импульс Берлаге задается аналитическим выражением:

$$g(t) = A \cdot t^{n(p_{max}) \cdot \Delta} \exp\left(-\frac{n(p_{max}) \cdot \Delta}{p_{max} \cdot t_{end}} \cdot t\right) \cdot \sin(2\pi ft),$$

где A — амплитуда, выбирается таким образом, чтобы $\|g(t)\| = 1$; t_{end} — длина атома; p_{max} — положение максимума относительно длины атома, $p_{max} \in [0.01, 0.4]$; f — частота от 200 до 20000 Гц; $n(p_{max})$ — предельное значение параметра n , вычисляемое по формуле:

$$n(p_{max}) = \frac{\ln 0.05}{\ln \frac{1}{p_{max}} - \frac{1}{p_{max}} + 1},$$

Δ — коэффициент варьирования параметра n относительно предельного значения.

Проведенные исследования показали, что импульсы Берлаге подходят для описания участков сигнала, содержащих импульсы ГАЭ, а импульсы Гаусса — для аппроксимации шумовой составляющей сигналов [13].

Задача разреженной аппроксимации (5) не разрешима за полиномиальное время. Алгоритм точного решения требует полного перебора всех возможных комбинаций функций из словаря, то есть имеет факториальную сложность $O(N!)$. Одним из наиболее часто используемых алгоритмов приближенного решения задачи разреженной аппроксимации является алгоритм согласованного преследования (Matching Pursuit, MP). Данный алгоритм был предложен Mallat и Zhang в 1993 году [21], он имеет квадратично-логарифмическую вычислительную сложность $O(N^2 \log N)$ (если для вычисления ковариации используется быстрое преобразование Фурье, в противном случае — кубическую вычислительную сложность $O(N^3)$). Алгоритм MP можно описать в виде итерационной процедуры:

$$\begin{cases} R_0(t) = s(t), \\ (m, h) = \arg \left[\max_{k,j} \langle g_k(t - \tau_j), R_i(t) \rangle \right], \\ R_{i+1}(t) = R_i(t) - \langle g_m(t - \tau_h), R_i(t) \rangle \cdot g_m(t - \tau_h), \end{cases}$$

где $R(t)$ — невязка, τ — сдвиг функции $g(t)$ относительно сигнала $s(t)$.

У алгоритма согласованного преследования имеется несколько существенных недостатков. Во-первых, для обеспечения достаточной точности разложений требуется использование словарей больших размеров, что в свою очередь вызывает степенной рост скорости выполнения алгоритма. Во-вторых, поскольку выбор атомов осуществляется из неизменяющегося словаря, полученные разложения отличаются «грубой» дискретизацией в пространстве параметров. Для решения вышеперечисленных проблем авторами было предложено улучшить классический алгоритм таким образом, чтобы на словарях ограниченного размера можно было строить разложения требуемой точности. Поскольку на каждой итерации алгоритма определяются параметры атома, имеющего наибольшее скалярное произведение с сигналом, то итерацию согласованного преследования можно описать в виде задачи поиска максимума функции многих переменных:

$$F(\tau, \mathbf{p}) = \langle g(t - \tau, \mathbf{p}), R(t) \rangle \rightarrow \max_{\mathbf{p}}$$

Главная идея предлагаемых улучшений — применять методы оптимизации для уточнения параметров \mathbf{p} функции, имеющей максимальное скалярное произведение с сигналом. Разработанный алгоритм

был назван адаптивное согласованное преследование (Adaptive Matching Pursuit, AMP) [13].

На основе изложенных выше соображений авторами были сформулированы уточненные требования к модели ГАЭ (1):

- функции $g_i(t)$ принадлежат линейному нормированному пространству $L^2(R)$;
- словарь $\mathbf{D} = \{g_i(t)\}_{1 \leq i \leq K}$ избыточен;
- $g_i(t)$ представляют собой сдвинутые по времени, модулированные функции Гаусса и Берлаге. Каждый атом однозначно определяется набором параметров \mathbf{p} : частотой f и параметрами, влияющими на форму огибающей импульса;
- атомы нормированы $\|g_i(t)\| = 1$.

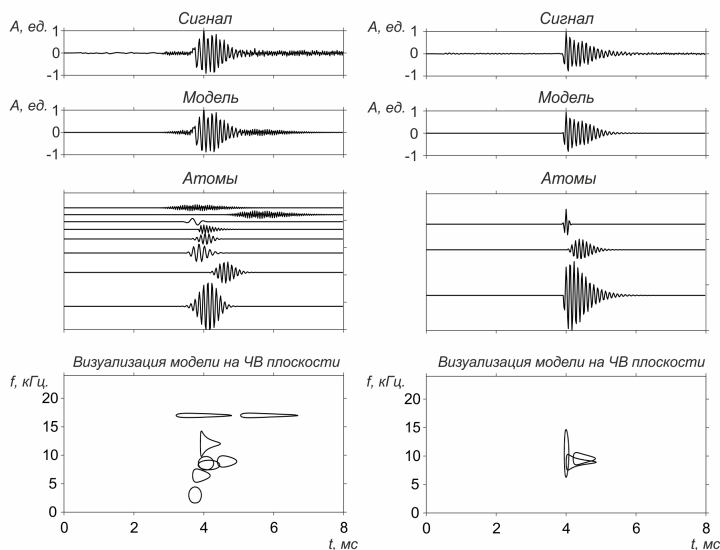


Рис. 7. Примеры моделей двух импульсов ГАЭ (контуры преобразования Вигнера—Вилля вычислены по уровню 0.2)

На рисунке 7 изображены модели двух импульсов ГАЭ. Количество атомов, обеспечивающее достаточную точность модели, определяется с помощью величины ERR , вычисляемой отношением нормы невязки к норме исходного сигнала:

$$ERR = \frac{\|R\|}{\|s\|} \cdot 100\%.$$

Разложение останавливается при достижении *ERR* порогового значения в 5%.

Для визуализации частотно-временной структуры моделей было выбрано преобразование Вигнера—Вилля, поскольку его частотно-временное разрешение минимально зависит от длительности анализируемого сигнала.

$$P_s(\tau, \omega) = \int_{-\infty}^{+\infty} s\left(\tau + \frac{t}{2}\right) \cdot s^*\left(\tau - \frac{t}{2}\right) \exp(-j\omega t) dt.$$

Однако поскольку преобразование нелинейно и строится на основе квадратичной функции от исходного сигнала, то для аддитивных сигналов в отображении проявляется сильная интерференционная компонента. Поэтому было решено вычислять преобразования Вигнера—Вилля отдельно для каждого из атомов, входящих в состав модели сигнала, и накладывать их отображения на частотно-временной плоскости (ЧВ плоскости) друг на друга.

5. Обсуждение результатов. Изложенная методика была реализована в прикладной программе [22] и применена для обработки и анализа данных наблюдений геоакустической эмиссии ИКИР ДВО РАН. Записи сигналов ГАЭ представлены в виде 15-минутных звуковых файлов. В ходе предварительной обработки каждого файла выделяются символы алфавита и ранжируются по значениям их размерности. Пример графического окна программы с результатами предобработки представлен на рисунке 8. Высота столбца отражает нормированное значение частоты появления каждого символа на интервале наблюдения в логарифмическом масштабе. Размерность символов обозначена цифрами справа сверху каждого столбца.

На рисунке 9 представлена динамика количественного состава множеств обнаруженных импульсов ГАЭ, а также множеств выделенных символов алфавитов в сигналах ГАЭ, зарегистрированных на станции «Карымшина» 01.01.2018. По рисунку 9а видно, что количество импульсов ГАЭ и символов алфавита, которые были выявлены в ходе предобработки 40 файлов, соответствующих временному периоду с 02:00:00 до 12:00:00, существенно не изменялось. В то время как на рисунке 9б видны довольно резкие изменения количества соответствующих параметров (временной период с 12:00:00 до 22:00:00). Обнаружение подобных аномалий вероятнее всего характеризует изменение условий генерации импульсов

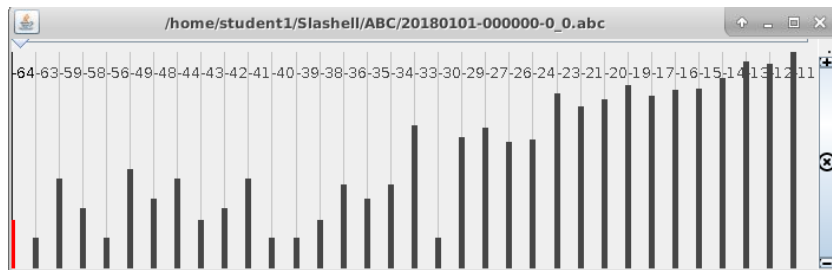


Рис. 8. Ранговое распределения символов алфавита 15-минутной записи сигнала ГАЭ

ГАЭ в среде, а, следовательно, может быть использовано как показатель изменения состояния среды под действием внешних факторов.

Предлагаемая методика кодирования импульсов с последующим выделением алфавита на выбранном интервале наблюдений позволяет рассчитать информационные показатели сигнала ГАЭ. В частности, оценки вероятности появления символов в выделенном алфавите и вычисленные на их основе показатели информационной энтропии. Для этого предусмотрена функция расчета частной энтропии каждого из символов выделенного алфавита по результатам оценки вероятности его появления согласно формуле (6), а также рассчитывается значение информационной энтропии для каждого алфавита по формуле (7):

$$H_i = -\log_2(p_i); \quad (6)$$

$$H(x) = -\sum_{i=1}^n p_i \log_2(p_i), \quad (7)$$

где H_i — частная энтропия для i -го символа алфавита; $p_i = m_i/M$ — оценка вероятности появления i -го символа; m_i — количество обнаруженных импульсов, описываемых паттерном символа i ; M — общее число импульсов, закодированных в n -символов ($\sum p_i = 1$); $H(x)$ — информационная двоичная энтропия для x независимых событий обнаружения импульсов, преобразуемых в n символов.

Частная энтропия (6) позволяет оценить информационный вклад (значимость) каждого символа, входящего в алфавит. Значение же информационной энтропии (7) отражает степень непредсказуемости состава выделяемого алфавита.

Указанные энтропийные показатели в настоящее время применяются для поиска закономерностей в поведении локальной системы

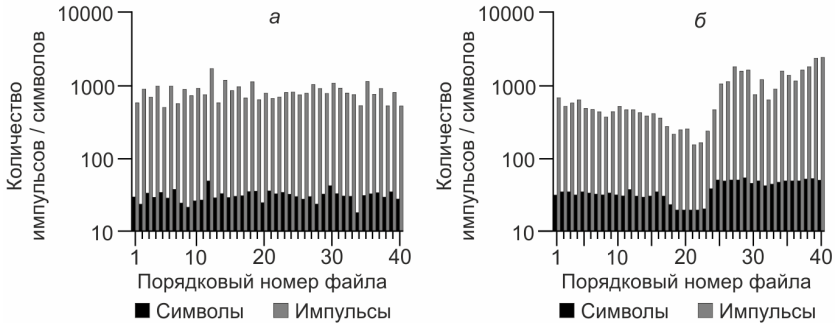


Рис. 9. Результаты преобработки сигнала ГАЭ: (а) — нормальная и (б) — anomальная картина динамики обнаруженного количества импульсов и количества символов

приповерхностных пород в зависимости от влияния внешних факторов. На рисунке 10 представлены примеры нормальной (а) и anomальной (б) картин динамики значений информационной энтропии, рассчитанной для последовательных 15-минутных интервалов сигнала ГАЭ в указанные выше временные периоды (с 02:00:00 до 12:00:00 (рисунок 10а) и с 12:00:00 до 22:00:00 (рисунок 10б) 01.01.2018). Anomalia указывает на заметное понижение неопределенности состава алфавита, связанное с некоторыми внутренними процессами в приповерхностных породах в районе измерений. Интересен тот факт, что картины anomалий (рисунки 9б и 10б) по-разному отражают наблюдаемую динамику. Данное явление еще не нашло своего объяснения.

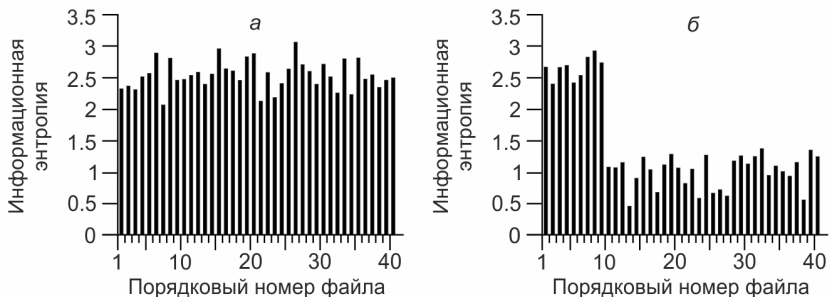


Рис. 10. Результаты преобработки сигнала ГАЭ: (а) — нормальная и (б) — anomальная картины динамики значений информационной энтропии

Таким образом, можно полагать, что представленные здесь показатели рангового распределения символов алфавита, динамика оценки числа выделяемых импульсов и размерности алфавитов, а также энтропия сигнала являются чувствительным инструментом проведения исследований изменчивости состояния системы приповерхностных слоев литосферы под действием внешних факторов в районе наблюдений, в том числе землетрясений, с использованием акустических средств регистрации.

6. Заключение. Результаты проведенных исследований показали, что представленная методика позволяет:

- выделять информативные признаки в сигналах геоакустической эмиссии по их качественным особенностям, а в перспективе и по качественно-количественным особенностям;
- определять подобные по структуре паттерны импульсных сигналов ГАЭ, сокращая тем самым размерность множества выделенных импульсов, порожденных множеством источников, а также интерференцией принимаемых сигналов;
- эффективно решать проблему обработки большого потока данных путем унификации описания импульсов и их систематизации.

Кроме того, на базе методики с применением информационного подхода к описанию сигналов удалось создать инструменты обработки и анализа сигналов ГАЭ, позволяющие проводить исследования по контролю состояния приповерхностных слоев литосферы под воздействием внешних факторов. Также разработанная методика успешно применяется для обработки и анализа импульсных электромагнитных сигналов. Таким образом, практика применения настоящей методики показала возможность ее использования при проведении исследований любых потоков импульсных сигналов. Дальнейшие исследования коллектива авторов предполагают разработку методов выявления закономерностей в поведении геоакустических и электромагнитных сигналов, в том числе и перед землетрясениями.

Литература

1. *Gregori G.P., Poscolieri M., Paparo G. et al.* “Storms of crustal stress” and AE earthquake precursors // *Natural Hazards and Earth System Sciences*. 2010. no. 10. pp. 319–337.
2. *Салтыков В.А.* О возможности использования приливной модуляции сейсмических шумов в целях прогноза землетрясений // *Физика Земли*. 2017. №2. С. 84–96.
3. *Ларионов И.А., Маратулец Ю.В, Мищенко М.А. и др.* Исследования акустической эмиссии приповерхностных осадочных пород на Камчатке // *Геосистемы переходных зон*. 2017. Т. 1, №3. С. 57–63.
4. *Giovanni G.P., Paparo G., Poscolieri M. et al.* The seismic sequence in Central Italy (August–November 2016) Acoustic Emission (AE) monitoring and analysis // *New Concepts in Global Tectonics Journal*. 2016. vol. 4. no. 4. pp. 637–663.

5. *Купцов А.В.* Изменение характера геоакустической эмиссии в связи с землетрясением на Камчатке // *Физика Земли*. 2005. №10. С. 59–65.
6. *Купцов А.В., Ларионов И.А., Шевцов Б.М.* Особенности геоакустической эмиссии при подготовке камчатских землетрясений // *Вулканология и сейсмология*. 2005. №5. С. 45–58.
7. *Гордиенко В.А., Гордиенко Т.В., Краснописцев Н.В. и др.* Аномалия высокочастотных сигналов геоакустической эмиссии как оперативный предвестник землетрясения // *Акустический журнал*. 2008. Т. 54, №1. С. 97–109.
8. *Маратулец Ю.В., Щербина А.О.* Методы исследования пространственной анизотропии геоакустической эмиссии // *Электронный журнал «Техническая акустика»*. 2008. URL: <http://ejta.org/ru/marapuletz>. (дата обращения: 11.03.2019).
9. *Пережогин А.С., Шевцов Б.М.* Модели напряженно-деформированного состояния горных пород при подготовке землетрясений и их связь с геоакустическими наблюдениями // *Информационные технологии*. 2009. Т. 14. №3. С. 48–57.
10. *Biot M.A.* Theory of propagation of elastic waves in fluid-saturated porous solid. Path II. Higher frequency range // *J. Acoust. Soc. Amer.* 1956. vol. 28. no. 2. pp. 179–191.
11. *Френкель Я.И.* К теории сейсмических и сейсмoeлектрических явлений во влажной почве // *Известия АН СССР. Серия география и геофизика*. 1944. Т. 8. №4. С. 133–149.
12. *Gassmann F.* Elastic waves through a parking of spheres // *Geophysics*. 1951. vol. 16. pp. 673–685.
13. *Маратулец Ю.В., Луковенкова О.О., Тристанов А.Б., Ким А.А.* Методы регистрации и частотно-временного анализа сигналов геоакустической эмиссии. Владивосток: Дальнаука, 2017. 148 с.
14. *Shcherbina A.O.* On some features of geoacoustic emission signals before earthquakes // *E3S Web Conferences*. 2017. vol. 20. 6 p. URL: https://www.e3s-conferences.org/articles/e3sconf/abs/2017/08/e3sconf_strpep2017_03005e3sconf_strpep2017_03005.html. (дата обращения: 12.03.2019).
15. *Сычев В.Н., Имашев С.А.* Оценка параметра Херста сейсмического сигнала // *Геосистемы переходных зон*. 2017. №2. С. 50–61.
16. *Мищенко М.А.* Статистика возникновения предсейсмических аномалий в геоакустической эмиссии и атмосферном электрическом поле // *Вестник КРАУНЦ. Физико-математические науки*. 2016. №3(14). С. 47–52.
17. *Senkevich Yu.I., Duke V.A., Mishchenko M.A., Solodchuk A.A.* Information approach to the analysis of acoustic and electromagnetic signals // *E3S Web of Conferences*. 2017. vol. 20. 02012. 9 p.
18. *Yang B.Y., Liu R.N., Chen X.F.* Sparse Time-Frequency Representation for Incipient Fault Diagnosis of Wind Turbine Drive Train // *IEEE Transactions on Instrumentation and Measurement*. 2018. vol. 67. no. 11. pp. 2616–2627.
19. *Raj S., Ray K.C.* Sparse representation of ECG signals for automated recognition of cardiac arrhythmias // *Expert systems with applications*. 2018. vol. 105. pp. 49–64.
20. *Huai S., Zhang S.* A novel sparse representation algorithm for AIS real-time signals // *Eurasip journal on wireless communications and networking*. 2018. URL: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1244-9>. (дата обращения: 11.03.2019).
21. *Mallat S.G., Zhang Z.* Matching pursuits with time-frequency dictionaries // *IEEE Transactions Signal Processing*. 1993. vol. 41. pp. 3397–3415.
22. *Сенкевич Ю.И.* Структурно-лингвистическая обработка геофизических сигналов и рядов данных // *Свидетельство о государственной регистрации программы для ЭВМ №2019617637*. 18.06.2019.

Сенкевич Юрий Игоревич — д-р техн. наук., доцент, ведущий научный сотрудник лаборатория акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН. Область научных интересов: прикладные аспекты цифровой обработки сигналов, разработка методов обработки и анализа нестационарных сигналов. Число научных публикаций — 70. senkevich@ikir.ru, www.ikir.ru; ул. Мирная, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Марапулец Юрий Валентинович — д-р физ.-мат. наук., доцент, директор, Институт космофизических исследований и распространения радиоволн ДВО РАН; профессор, кафедра информатики физико-математического факультета, Камчатский государственный университет им. Витуса Беринга(КамГУ); профессор, кафедра систем управления факультета информационных технологий, Камчатский государственный технический университет(КамчатГТУ). Область научных интересов: цифровая обработка сигналов, физическая акустика, геоакустика. Число научных публикаций — 122. marpl@ikir.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Луковенкова Ольга Олеговна — канд. техн. наук., старший научный сотрудник, лаборатория акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН, доцент, кафедра информатики физико-математического факультета, Камчатский государственный университет им. Витуса Беринга(КамГУ), доцент, кафедра систем управления факультета информационных технологий, Камчатский государственный технический университет(КамчатГТУ). Область научных интересов: цифровая обработка сигналов, интеллектуальный анализ временных рядов. Число научных публикаций — 34. o.o.lukovenkova@yandex.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Солодчук Александра Андреевна — канд. физ.-мат. наук., научный сотрудник, лаборатория акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН. Область научных интересов: физическая акустика, геоакустика, разработка методов анализа природных сигналов. Число научных публикаций — 42. aleksandra@ikir.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Поддержка исследований. Работа выполнена при финансовой поддержке Российского научного фонда (проект № 18-11-00087).

Yu.I. SENKEVICH, Yu.V. MARAPULETS, O.O. LUKOVENKOVA,
A.A. SOLODCHUK

TECHNIQUE OF INFORMATIVE FEATURES SELECTION IN GEOACOUSTIC EMISSION SIGNALS

Senkevich Yu.I., Marapulets Yu.V., Lukovenkova O.O., Solodchuk A.A. Technique of Informative Features Selection in Geoaoustic Emission Signals.

Abstract. Studies of geoaoustic emission in a seismically active region in Kamchatka show that geoaoustic signals produce pronounced pulse anomalies during the earthquake preparation and post-seismic relaxation of the local stresses field at the observation point. The qualitative selection of such anomalies is complicated by a strong distortion and weakening of the signal amplitude. A review of existing acoustic emission analysis methods shows that most often researchers turn to the analysis of more accessible to study statistical properties and energy of signals. The distinctive features of the approach proposed by the authors are the extraction of informative features based on the analysis of time and frequency-time structures of geoaoustic signals and the description of various forms of recognizable pulses by a limited pattern set. This study opens up new ideas to develop methods for detecting anomalous behavior of geoaoustic signals, including anomalies before earthquakes.

The paper describes a technique of information extraction from geoaoustic emission pulse streams of sound frequency range. A geoaoustic pulse mathematical model, representing the signal generation process from a variety of elementary sources, is presented. A solution to the problem of detection of geoaoustic signal informative features is presented by the means of description of signal fragments by the matrixes of local extrema amplitude ratios and of interval ratios between them. The result of applying the developed algorithm to describe automatically the structure of the detected pulses and to form a pattern set is shown. The patterns characterize the features of geoaoustic emission signals observed at IKIR FEB RAS field stations. A technique of reduction of the detected pulse set dimensions is presented. It allows us to find patterns similar in structure. A solution to the problem of processing of a large data flow by unifying pulses description and their systematisation is proposed. The results of the research allowed the authors to create a tool to investigate the dynamic properties of geoaoustic emission signal in order to develop earthquake prediction detectors.

Keywords: Geoaoustic Emission, Geoaoustic Pulse Model, Sparse Approximation, Informative Features, Signal Patterns.

Senkevich Yuri Igorevich — Ph.D., Dr.Sci., Associate Professor, Leading Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS. Research interests are applied aspects of digital signal processing, development of processing and analysis methods of non-stationary signals. The number of publications is 70. senkevich@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Marapulets Yuri Valentinovich — Ph.D., Dr.Sci., Associate Professor; Director, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS; Professor, Department of Informatics of Faculty of Physics and Mathematics, KamGU Vitus Bering; Professor, Department of control systems of Faculty of Information Technology, KamchatSTU. Research interests are digital signal processing, physical acoustics, geoaoustics. The number of publications is 122.

marpl@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Lukovenkova Olga Olegovna — Ph.D., Senior Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS; Associate Professor; Department of Informatics of Faculty of Physics and Mathematics, KamGU Vitus Bering; Associate Professor, Department of Control Systems of Faculty of Information Technology, KamchatSTU. Research interests are digital signal processing, time series data mining. The number of publications is 34. o.o.lukovenkova@yandex.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Solodchuk Alexandra Andreevna — Ph.D., Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS. Research interests are physical acoustics, geoaoustics, development of analysis methods of natural signals. The number of publications is 42. aleksandra@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Acknowledgements. This research is supported by Russian Science Foundation (project 18-11-00087).

References

1. Gregori G.P., Poscolieri M., Paparo G. et al. [“Storms of crustal stress” and AE earthquake precursors]. *Natural Hazards and Earth System Sciences*. 2010. no. 10. pp. 319–337.
2. Saltykov V.A. [On the possibility of using the tidal modulation of seismic waves for forecasting earthquakes]. *Izvestiya. Physics of the Solid Earth*. 2017. vol. 53. no. 2. pp. 250–261.
3. Larionov I.A., Marapulets Yu.V., Mishchenko M.A. et al. [Issledovaniia akusticheskoi emissii pripoverkhnostnykh osadochnykh porod na Kamchatke]. *Geosistemy perekhodnykh zon — Geosystems of Transition Zones*. 2017. vol. 1, no. 3. pp. 57–63. (In Russ.).
4. Giovanni G.P., Paparo G., Poscolieri M. et al. [The seismic sequence in Central Italy (August–November 2016) Acoustic Emission (AE) monitoring and analysis]. *New Concepts in Global Tectonics Journal*. 2016. vol. 4. no. 4. pp. 637–663.
5. Kuptsov A.V. [Variations in the geoaoustic emission pattern related to earthquakes on Kamchatka]. *Izvestiya. Physics of the Solid Earth*. 2005. vol. 41. no. 10. pp. 825–831.
6. Kuptsov A.V., Larionov I.A., Shevtsov B.M. [Osobennosti geoakusticheskoi emissii pri podgotovke kamchatskikh zemletriasenii]. *Vulkanologiya i seismologiya — Volcanology & Seismology*. 2005. no. 5. pp. 45–58. (In Russ.).
7. Gordienko V.A., Gordienko T.V., Krasnopistsev N.V. et al. [Anomaly in high-frequency geoaoustic emission as a close earthquake precursor]. *Acoustical physics*. 2008. vol. 54, no. 1. pp. 82–93.
8. Marapulets Yu.V., Shcherbina A.O. [Metody issledovaniia prostranstvennoi anizotropii geoakusticheskoi emissii]. *Elektronnyi zhurnal “Tekhnicheskaya akustika” — Electronic Journal Technical Acoustics*. 2008. Available at: <http://ejta.org/ru/marapuletz>. (accessed 11.03.2019). (In Russ.).
9. Perezhogin A.S., Shevtsov B.M. [Modeli napriazhenno-deformirovannogo sostoianiia gornyykh porod pri podgotovke zemletriasenii i ikh svyaz s geoakusticheskimi nabliudeniiami]. *Vychislitelnye tekhnologii — Computational Technologies*. 2009. vol. 14, no. 3. pp. 48–57. (In Russ.).
10. Biot M.A. [Theory of propagation of elastic waves in fluid-saturated porous solid. Path II. Higher frequency range]. *J. Acoust. Soc. Amer*. 1956. vol. 28. no. 2. pp. 179–191.

11. Frenkel Ya.I. [K teorii seismicheskikh i seismoelektricheskikh iavlenii vo vlazhnoi pochve]. *Izvestia AN SSSR. Seriya geografii i geofizika*. 1944. vol. 8. no. 4. pp. 133–149. (In Russ.).
12. Gassmann F. [Elastic waves through a parking of spheres]. *Geophysics*. 1951. vol. 16. pp. 673–685.
13. Marapulets Iu.V., Lukovenkova O.O., Tristanov A.B., Kim A.A. *Metody registratsii i chastotno-vremennogo analiza signalov geoakusticheskoi emissii* [Methods for recording and for time-frequency analysis of geoacoustic emission signals]. Vladivostok: Dalnauka, 2017. 148 p. (In Russ.)
14. Shcherbina A.O. [On some features of geoacoustic emission signals before earthquakes]. *E3S Web Conferences*. 2017. vol. 20. 6 p. Available at: https://www.e3s-conferences.org/articles/e3sconf/abs/2017/08/e3sconf_strpep2017_03005/e3sconf_strpep2017_03005.html. (accessed 12.03.2019).
15. Sychev V.N., Imashev S.A. [Otsenka parametra Khersta seismicheskogo signala]. *Geosistemy perekhodnykh zon — Geosystems of Transition Zones*. 2017. no. 2. pp. 50–61.
16. Mishchenko M.A. [Statistika vozniknoveniia predeismicheskikh anomalii v geoakusticheskoi emissii i atmosfernom elektricheskom pole]. *Vestnik KRAUNTs. Fiziko-matematicheskie nauki — Bulletin KRASEC. Physical & Mathematical Sciences*. 2016. no. 3(14). pp. 47–52. (In Russ.).
17. Senkevich Yu.I., Duke V.A., Mishchenko M.A., Solodchuk A.A. Information approach to the analysis of acoustic and electromagnetic signals. *E3S Web of Conferences*. 2017. vol. 20. 02012. 9 p.
18. Yang B.Y., Liu R.N., Chen X.F. [Sparse Time-Frequency Representation for Incipient Fault Diagnosis of Wind Turbine Drive Train]. *IEEE Transactions on Instrumentation and Measurement*. 2018. vol. 67. no. 11. pp. 2616–2627.
19. Raj S., Ray K.C. [Sparse representation of ECG signals for automated recognition of cardiac arrhythmias]. *Expert systems with applications*. 2018. vol. 105. pp. 49–64.
20. Huai S., Zhang S. [A novel sparse representation algorithm for AIS real-time signals]. *Eurasip journal on wireless communications and networking*. 2018. Available at: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1244-9>. (accessed 11.03.2019).
21. Mallat S.G., Zhang Z. [Matching pursuits with time-frequency dictionaries]. *IEEE Transactions Signal Processing*. 1993. vol. 41. pp. 3397–3415.
22. Senkevich Yu.I. [Structural and linguistic processing of geophysical signals and data series]. Certificate of state registration of computer programs №2019617637. 18.06.2019. (In Russ.)

К.А. БАТЕНКОВ

ТОЧНЫЕ И ГРАНИЧНЫЕ ОЦЕНКИ ВЕРОЯТНОСТЕЙ СВЯЗНОСТИ СЕТЕЙ СВЯЗИ НА ОСНОВЕ МЕТОДА ПОЛНОГО ПЕРЕБОРА ТИПОВЫХ СОСТОЯНИЙ

Батенков К.А. Точные и граничные оценки вероятностей связности сетей связи на основе метода полного перебора типовых состояний.

Аннотация. В работе рассматривается один из методов анализа и синтеза структур сетей связи, основанный на наиболее простом подходе к вопросу расчета вероятности связности — методе полного перебора типовых состояний сети. При этом под типовыми состояниями сети понимаются события связности и несвязности графа сети, представляющие собой простые цепи и сечения данного графа. Несмотря на существенный недостаток метода полного перебора типовых состояний, который заключается в значительной трудоемкости проводимых вычислений, он оказывается достаточно востребованным. Кроме того, на его основе возможно получать граничные оценки вероятности связности сети. Так, при расчете границ Эзари — Прошана используется полный набор несвязных (для верхней) и связных (для нижней) состояний сети связи. Данные границы основаны на утверждении, что вероятность связности сети при тех же условиях выше (ниже), чем у сети, составленной из последовательного (параллельного) соединения полного набора независимых несвязных (связных) подграфов. При расчете границ Литвака — Ушакова используются только реберно-непересекающиеся сечения (для верхней) и связные подграфы (для нижней), то есть подмножества элементов такие, в которых какой-либо элемент не встречается дважды. В данной границе учтено широко известное естественное свойство монотонности, заключающееся в уменьшении (увеличении) надежности сети при снижении (повышении) надежности любого элемента. С точки зрения сложности вычислительных процедур границы Эзари — Прошана имеют существенный недостаток: они предполагают определение всех связных подграфов для расчета верхней границы и минимальных разрезов для нижней, что само по себе нетривиально. Границы Литвака — Ушакова подобными недостатками не страдают: вычисляя их, можно ограничиться перебором необходимого числа вариантов наборов независимых связных и несвязных состояний графа.

Ключевые слова: сеть связи, граф, структура, вероятность связности, метод полного перебора типовых состояний, граничные оценки.

1. Введение. Сложные многокомпонентные сети являются характерной чертой современного общества [1]. Урбанизация и глобализация, сопровождающиеся быстрым развитием инфраструктуры и технологий, создали возможности построения производительных сетей со все более взаимозависимыми компонентами. Данные сети предназначены для распределения какого-либо ресурса или товара, а их наиболее яркими представителями являются сети связи, например телефонные сети или Интернет. Кроме того, к ним относятся транспортные сети (сети автомобильных, железных дорог и авиасообщений), инженерные сети (сети электро-, газо- и водоснабжения) и так далее.

Системы в виде сетей — неотъемлемая часть мегаполисов; их поломки, хотя и редкие, имеют серьезные последствия для обществен-

ной безопасности [2]. Зависимость нашего общества от технологических систем и информационных сетей растет. В настоящее время мы едва ли можем представить себе нашу жизнь без электричества, телефонов, самолетов и даже без Интернета. В результате возникает потребность в высокой устойчивости современных сетей. При этом количественная оценка устойчивости (и как ее составляющие надежности и живучести) подобных сетей является значимым аспектом проектирования, оптимизации и эксплуатации сети [3, 4].

Если остановиться конкретно на телекоммуникационных сетях, то в последние годы наблюдается резкий всплеск объемов циркулирующего трафика. Ощутим соответствующий рост спроса на компьютерные и коммуникационные сети как по сложности, так и по размерам. По мере роста размеров сетей одновременно варьируется в сложно прогнозируемых пределах количество и частота отказов компонентов. Поэтому целесообразно проектировать сети, способные сохранять работоспособность, несмотря на выход из строя оборудования.

Целью настоящей работы является анализ точных и граничных оценок устойчивости (надежности и живучести) сложных многокомпонентных сетей на основе метода полного перебора типовых состояний, как связанных, так и несвязных. При этом рассматриваются два крайних случая события связности сети — наличие хотя бы одного простого пути между заданной парой вершин и остового дерева, связывающего все вершины между собой, и события несвязности — присутствие минимальных сечений как для направления связи, так и для графа в целом, то есть между любой парой вершин. Исходными данными для анализа сетей выступают вероятности связности их отдельных элементов (ребер графа), представляющие собой аналитический эквивалент коэффициента готовности (надежность) или оперативной готовности (живучесть), на основе которых определяются значения вероятности связности сети в целом.

В зависимости от специфики сетей, их типа (вида) следует оценивать структурную надежность сетей с помощью различных показателей, подразделяемых на детерминированные и вероятностные. Первые из них зависят от топологии сети и определяют степень сложности для вывода ее из строя. В качестве стохастической меры надежности сети используется вероятность связности. Еще раз подчеркнем, что здесь и далее рассматриваются показатели надежности сетей, которые достаточно просто могут быть преобразованы в показатели живучести.

Все методы анализа показателей устойчивости (надежности и живучести) сетей связи можно разделить на две крупные категории: точные и приближенные. Что же касается различий с точки зрения

используемых показателей, то методы также условно делятся на два класса: детерминированные и вероятностные, причем если для детерминированных методов характерно использование в каждом методе своего собственного показателя, то в стохастических он всего один (все оставшиеся являются производными от исходного) — вероятность связности или несвязности. Отличаются же последние друг от друга исключительно используемым математическим аппаратом [4].

Основным недостатком детерминированных методов анализа является сложность их сравнительного анализа в виду различной физической природы используемых показателей. В результате наиболее часто используемыми оказываются вероятностные методы, различие между которыми связано в основном лишь в способе преобразования понятия связности сети, формализуемого в форме различного рода логических функций, в вероятность связности. При этом наиболее наглядными и естественными выглядят методы перебора типовых состояний, поскольку достаточно просто проецируют понятия связного и несвязного состояний на математический язык, а также используют весьма простые способы преобразования логических функций в алгебраическую форму, позволяющую с помощью операций сложения и умножения рассчитывать требуемый показатель.

2. Метод полного перебора связных состояний сети связи.

Следует отметить, что в научной литературе, посвященной анализу надежности сетей связи, метод полного перебора связных состояний называется методом прямого перебора состояний путей системы связи [4] или методом прямого перебора простых цепей [5], а в источниках классической теории надежности — методом минимальных путей [6]. Однако поскольку необходимо проанализировать два случая графов сетей связи — многосвязные и двухсвязные, то целесообразнее рассматривать состояние связности графа как общее понятие, объединяющее и простые цепи, и остовые деревья.

Событию связности многополюсной или двухполюсной сети соответствует факт существования хотя бы одного остового дерева либо простой цепи. Далее остовое дерево и простую цепь будем обозначать термином «связный подграф» [4]. Таким образом, одновременное присутствие двух и более связных подграфов вполне возможно, а значит, для расчета вероятности связности необходимо использовать теорему сложения для совместных событий [5], которая формулируется как совокупность логических высказываний типа «или», откуда следует применимость схемы «или то, или другое, или оба вместе». Следует заметить, что схема «либо» (исключающее «или») предполагает использование высказывания «либо то, либо другое».

Следует отметить, что в ГОСТ Р 53111–2008 [7] определено понятие *вероятность связности* как вероятности исправного состояния хотя бы одной цепи из всех возможных цепей или (при ограничении числа транзитных участков) хотя бы одной цепи с допустимым рангом. При этом при практических расчетах перечень простых цепей или путей между узлами ограничивают только теми путями, которые содержат допустимое число транзитных участков, зависящее от допустимого уровня искажений передаваемой по линии связи информации. В данной работе рассматривается более общий случай, дополнительно затрагивающий событие исправности хотя бы одного остового дерева.

Сущность данного метода расчета вероятности связности заключается в формировании всех существующих вариантов из n связных подграфов так, что каждая альтернатива включает все элементы графа, принадлежащие формируемому связным подграфам, а также расчету вероятности существования альтернатив и знакопеременному их суммированию [8]. Метод оказывается весьма трудоемким, так как предполагает рассмотрение 2^n всех возможных вариантов из n связных подграфов, что даже с использованием современных процессоров не всегда допустимо с точки зрения реализации для сложно разветвленных структур сетей связи. В то же время значительным достоинством метода является отсутствие необходимости вычисления индикатора (функции) связности для всех подграфов исходного графа сети [9].

Пусть $G_i, i = 1, 2, \dots, n$ — множество всех подграфов (остовов или простых цепей) исходного графа G . Событие, заключающееся в исправности всех элементов связного подграфа G_i , также обозначено G_i . Доказано [6], что множество всех исправных состояний сети связи

$\bigcup_{i=1}^n G_i$ совпадает с объединением событий G_i , а значит, для вероятности связности сети правомерно равенство [6]:

$$p_G = P\left(\bigcup_{i=1}^n G_i\right) = \sum_{i=1}^n P(G_i) - \sum_{i < j} P(G_i G_j) + \sum_{i < j < k} P(G_i G_j G_k) - \dots + (-1)^{n-1} P(G_1 G_2 \dots G_n). \quad (1)$$

где $P(G_1 G_2 \dots G_i)$ — вероятность пересечения событий G_1, G_2, \dots, G_i .

Учитывая, что вероятность совместного события связности сразу нескольких подграфов рассчитывается на основе соответствующих

условных вероятностей, формулу вероятности связности сети (1) можно представить в виде:

$$p_G = \sum_{i=1}^n p_i - \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_{j,i} + \\ + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n p_i p_{j,i} p_{k,j,i} - \dots + (-1)^{n-1} p_1 p_{2,1} \cdot \dots \cdot p_{n,n-1, \dots, 2,1},$$

где $p_i = P(G_i)$ — вероятность связности подграфа G_i , $i = 1, 2, \dots, n$; $p_{j,j-1, \dots, i+1, i} = P(G_i / G_{i+1} \dots G_{j-1}, G_j)$ — вероятность связности подграфа G_i , $i = 1, 2, \dots, n$ без учета элементов, входящих в подграфы G_{i+1} , ..., G_{j-1} , G_j .

Следует отметить, что по своей сути процедура вычисления вероятности связности подграфа G_i , $i = 1, 2, \dots, n$ без учета элементов, входящих в подграфы G_{i+1} , ..., G_{j-1} , G_j , эквивалентна расчету вероятности связности совокупного подграфа, который состоит из элементов графа, входящих хотя бы в одну из объединяемых подграфов в данном слагаемом [5].

Таким образом, процесс накопления точного значения вероятности связности p_G является аддитивным и знакопеременным [5]. Следует заметить, что результаты суммирований при количестве сумм, меньших числа связных подграфов, в общем случае могут выходить за пределы допустимых значений вероятности, то есть за интервал $[0, 1]$. При этом значения нечетных слагаемых всегда будут больше значений четных. Следует также отметить особенность данного метода, которая состоит в том, что чем больше значения вероятностей исправного состояния элементов графа, тем «точнее» слагаемые повторяют поведение огибающих биномиальных коэффициентов.

3. Методы формирования множеств связных подграфов. Сами по себе процедуры формирования множеств как простых цепей, так и остовых деревьев являются достаточно нетривиальными. В первую очередь это связано с необходимостью накопления достаточно большого и зачастую частично дублирующего самих себя наборов связных подграфов, количество которых в зависимости от разветвленности исходной сети может достигать сотен, тысяч и даже миллионов комбинаций.

В основе метода формирования множества простых цепей находится предположение о том, что данное множество состоит из всех

возможных комбинаций последовательного перебора связанных друг с другом ребер графа, за исключением тех комбинаций, которые формируют циклы. В связи с этим длина любой приемлемой комбинации не может превышать числа вершин в рассматриваемом графе, а начальная и конечная вершины пути должны, соответственно, содержаться в первом и последнем ребре комбинации. Подобную процедуру достаточно просто реализовать с помощью существующих пакетов прикладных программ математического моделирования, например MathCad, используя в качестве исходных данных матрицу связностей графа сети и номера начальной и конечной вершин направления связи.

Так, первоначально проверяется на наличие единичных (ненулевых) элементов строки (столбца) матрицы смежностей, соответствующей вершине–источнику. Запоминаются все отобранные вершины и для каждой проверяется условие, является ли найденная вершина стоком или нет. При удовлетворении данного условия этот путь сохраняется, а при не выполнении — проводится дальнейший поиск по строкам (столбцам), соответствующим всем оставшимся вершинам, которые не являются стоками, с последующей проверкой вышеупомянутого условия. Для каждого из сохраняемого пути проводится дополнительная проверка на наличие в нем петель (циклов) и его сохранение осуществляется только при их отсутствии. Поиск производится до тех пор, пока либо не будут перебраны все комбинации, то есть они приведут к стокам, либо количество ребер в маршрутах не будет на единицу меньше числа вершин.

Отметим, что подобная процедура оказывается весьма простой не только с точки зрения программной реализации, но и с позиции вычислительной сложности.

В основе метода формирования множества остовых деревьев положено понятие декартового произведения множества сечений вершин [10]. Под сечением вершины при этом понимается множество инцидентных ей ребер. Само декартово произведение трактуется традиционно — производится последовательный перебор всех пар, один из элементов которой принадлежит первому множеству, а второй — второму. Результаты произведения, имеющие в своем составе повторения, уничтожаются. Число выполняемых произведений на две единицы меньше количества вершин в рассматриваемом графе. Отметим, что реализация данного метода также возможна с использованием программы MathCad, а в качестве исходных данных используется только матрица смежностей графа исходной сети.

Несмотря на достаточно простую формализацию, данная процедура довольно громоздка с точки зрения задействованной памяти. Так,

в дальнейшем будет показано, что данное обстоятельство связано, прежде всего, с существенно большим числом остовых деревьев в рассматриваемом графе по сравнению с присутствующими в нем простыми путями. Действительно любое остовое дерево содержит один простой путь, но один и тот же путь может содержаться сразу в нескольких остовых деревьях.

Область практического применения данного метода в основном совпадает с областью применения метода перебора состояний элементов. Он применяется также и для расчета вероятности связности при значительном числе элементов, но небольшом числе связных подграфов сети [11].

4. Метод полного перебора несвязных состояний сети связи.

Данный метод, по сути, является антиподом метода перебора связных состояний сети связи. Последовательность решения задач сохраняется, но рассматриваются события несвязности. При этом событию несвязности многополюсной или двухполюсной сети соответствует факт несуществования ни одного остового дерева либо простой цепи. В научной литературе по классической теории надежности метод полного перебора несвязных состояний сети связи является частной вариацией метода минимальных сечений [6]. Аналогично ранее рассмотренному подходу необходимо остановиться на двух случаях графов сетей связи — много-связных и двухсвязных. В связи с этим событие несвязности как для многополюсной, так и для двухполюсной сети наступает в случае отказа всех элементов некоторого сечения. При этом в многополюсной сети сечением будет набор элементов, приводящий к несвязному графу, а в двухполюсной — к отсутствию маршрутов между источником и стоком.

Одновременное присутствие двух и более сечений вполне возможно, а значит, аналогично предыдущим выводам для расчета вероятности несвязности необходимо использовать теорему сложения для совместных событий [5].

Сущность данного метода расчета вероятности несвязности заключается в формировании всех существующих вариантов из n сечений так, что каждая альтернатива включает все элементы графа, принадлежащие формируемому сечению, а также расчету вероятности существования альтернатив и знакопеременному их суммированию [7, 8]. Следует отметить, что трудоемкость данного метода сопоставима со сложностью метода перебора связных состояний, а требование к вычислению индикаторов (функций) связности для всех подграфов исходного графа сети также отсутствует [12].

Пусть G_i^j , $i = 1, 2, \dots, n$ — множество всех сечений (многополюсной или двухполюсной сети) исходного графа G . Событие, за-

ключающееся в неисправности всех элементов сечения G'_i , обозначено G'_i и далее трактуется как событие несвязности сечения G'_i . Отметим, что в отличие от события связности подграфа G_i , которое возникает при неисправности хотя бы одного из ребер подграфа G_i , событие несвязности сечения G'_i возникает только при неисправности всех ребер сечения G'_i . Аналогично расчету вероятности связности множество всех неисправных состояний сети связи $\bigcup_{i=1}^n G'_i$ совпадает с объединением событий G'_i , поэтому для вероятности несвязности сети справедливо равенство:

$$q_G = Q\left(\bigcup_{i=1}^n G'_i\right) = \sum_{i=1}^n Q(G'_i) - \sum_{i<j} Q(G'_i G'_j) + \sum_{i<j<k} Q(G'_i G'_j G'_k) - \dots + (-1)^{n-1} Q(G'_1 G'_2 \dots G'_n), \quad (2)$$

где $Q(G'_1 G'_2 \dots G'_i)$ — вероятность пересечения событий G'_1, G'_2, \dots, G'_i .

Учитывая, что вероятность совместного события несвязности сразу нескольких сечений рассчитывается на основе соответствующих условных вероятностей, формулу вероятности несвязности сети (2) можно представить в виде:

$$q_G = \sum_{i=1}^n q_i - \sum_{i=1}^{n-1} \sum_{j=i+1}^n q_i q_{j,i} + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^n q_i q_{j,i} q_{k,j,i} - \dots + (-1)^{n-1} q_1 q_{2,1} \dots q_{n,n-1,\dots,2,1},$$

где $q_i = 1 - p_i = Q(G'_i)$ — вероятность несвязности сечения G'_i , $i = 1, 2, \dots, n$, $i = 1, 2, \dots, n$; $q_{j,j-1,\dots,i+1,i} = P(G'_i / G'_{i+1} \dots G'_{j-1}, G'_j)$ — вероятность несвязности сечения G'_i , $i = 1, 2, \dots, n$ без учета элементов, входящих в сечения $G'_{i+1}, \dots, G'_{j-1}, G'_j$.

Отсюда видно, что процедура накопления точного значения вероятности несвязности q_G является, так же как и вероятности

связности, аддитивной и знакопеременной. При этом все вышеописанные свойства идентичны для методов на основе как связных, так и несвязных подграфов.

5. Методы формирования множеств несвязных подграфов.

Процедуры формирования множеств сечений как для направления связи, так и для графа в целом являются несколько более простыми, чем аналогичные процедуры для связных подграфов. Данное обстоятельство обусловлено, прежде всего, тем, что для большинства реальных сетевых топологий количество соответствующих связных подграфов (путей и остовых деревьев) обычно сопоставимо или существенно больше количества несвязных подграфов (сечений в направлении и в целом).

Множества сечений в направлении связи формируются на основе ранее использованных сечений вершин графа [10]. При этом данные сечения разбиваются на три группы. Первая содержит только сечение вершины–истока, вторая — только сечение вершины–стока, а третья — все оставшиеся. Производится последовательный перебор всех комбинаций третьей группы, порядок которых варьируется от одного до значения на три единицы меньшего числа вершин в графе. Для каждой комбинации проверяется два условия. Первое: если комбинация содержит все вершины, входящие в первую группу (сечение вершины–истока), то она отбрасывается. Второе: если комбинация не включает в себя сечений, которые не содержат вершины, входящие во вторую группу (сечение вершины–стока), то она также отбрасывается. Данные условия позволяют существенно сократить количество не минимальных сечений, но полностью, как показали исследования, эту проблему не устраняют.

Каждая из полученных комбинаций третьей группы дополняется первой (сечением вершины–истока) и после удаления дублированных ребер сохраняется как сечение графа в направлении. Следует заметить, что в отличие от подхода, описанного в [10], необходимо дополнительно проверять сохраняемые сечения на минимальность, так как данная процедура формирует сечения, способные потенциально содержать в себе минимальные. Отметим, что подобная процедура также достаточно просто реализуется в программной среде MathCad, а в качестве исходных данных используются, как и при формировании простых путей, матрица смежностей и две вершины — исток и сток.

Множества сечений для графа в целом формируются на основе перебора всех допустимых комбинаций ребер от первого порядка до порядка, в общем случае соответствующего числу ребер в графе сети. При этом для сокращения количества рассматриваемых комбинаций используется прием, когда комбинация, являющаяся сечением, в дальнейшем не

порождает комбинации более высокого порядка. Для каждой комбинации проверяется два условия. Первое: является ли она сечением, то есть на основе метода поиска в ширину осуществляется контроль связности графа, не включающего ребра данной комбинации. Второе: выполняется проверка на минимальность данной комбинации, то есть является ли она подмножеством уже найденных сечений или нет. Реализация данной процедуры также достаточна проста в пакете математического моделирования MathCad, а исходные данные аналогичны используемым при формировании множества остовых деревьев — матрица смежностей.

Отметим, что в отличие от метода формирования сечений в направлении связи этот метод оказывается более громоздким вследствие как несколько большего количества сечений, так и более разветвленной процедуры образования комбинаций для проверки.

6. Граничные оценки. Ранее были рассмотрены аналитические методы вычисления точного значения вероятности связности выбранной пары вершин для случая двухполюсной сети [13, 14]. При этом подчеркивалось, что вычисление точного значения вероятности связности на сложно разветвленных сетях из-за большого объема трудозатрат не всегда реализуемо [15]. Кроме того, при проведении исследований структурной надежности сетей связи проектировщика интересует относительная точность вычислений, а не абсолютная, например, точность в первом, втором, третьем и так далее знаке (или значащей цифре) после запятой в значении вероятности связности [5, 16]. Однако эта точность должна быть гарантированной [17]. В этом случае следует оперировать термином «граничная оценка» вероятности связности, оговаривая точность (достигнутую или заданную) проведенных вычислений.

Следует отметить еще одно важное достоинство рассматриваемых методов перебора типовых (связных или несвязных) состояний — на их основе оказывается возможным получать граничные оценки вероятности связности сети. Так, при расчете границ Эзари — Прошана [18] используется полный набор несвязных (для верхней) и связных (для нижней) состояний сети связи. Данные границы основаны на утверждении, что вероятность связности сети при тех же условиях выше (ниже), чем у сети, составленной из последовательного (параллельного) соединения полного набора независимых несвязных (связных) подграфов. При расчете границ Литвака — Ушакова используются только реберно-непересекающиеся сечения (для верхней) и связные подграфы (для нижней), то есть подмножества элементов такие, в которых какой-либо элемент не встречается дважды. В данной границе учтено широко известное естественное свойство монотонности, заключающееся в уменьшении (увеличении) надежности сети при снижении (повышении) надежности любого элемента.

7. Верхняя граница Эзари — Прошана. При расчете верхней границы Эзари — Прошана используется полный набор связанных состояний сети связи (простых цепей).

Напомним, что $G_i, i = 1, 2, \dots, n$ — множество всех подграфов (простых цепей) исходного графа G , а событие, заключающееся в исправности всех элементов связанного подграфа G_i , также обозначено

G_i . Кроме того, множество всех исправных состояний сети связи $\bigcup_{i=1}^n G_i$ совпадает с объединением событий G_i , а значит, для вероятности связности сети на основе правила де Моргана [19] справедливо равенство:

$$p_G = P\left(\bigcup_{i=1}^n G_i\right) = P\left(\overline{\bigcap_{i=1}^n \overline{G_i}}\right),$$

где \bar{x} — событие противоположное событию x .

Следовательно:

$$q_G = 1 - p_G = P\left(\bigcap_{i=1}^n \overline{G_i}\right).$$

Вероятность совместного осуществления связанных случайных событий больше, чем произведение вероятностей этих событий. Значит, справедливо неравенство:

$$q_G \geq \prod_{i=1}^n P(\overline{G_i}).$$

Данное выражение показывает, что вероятность несвязности сети больше, чем вероятность неработоспособности параллельно соединенных всех независимых минимальных связанных подграфов. Подобное утверждение равносильно тому, что вероятность связности сети меньше при прочих равных условиях, чем у сети, включающей параллельное соединение всего набора независимых связанных подграфов. Таким образом, верхняя оценка Эзари — Прошана для вероятности связности сети имеет вид [18]:

$$p_G \leq p_e^+ = 1 - \prod_{i=1}^n q_i,$$

где $q_i = P(\overline{G_i})$ — вероятность несвязности связанного подграфа $G_i, i = 1, 2, \dots, n$.

Если связный подграф G_i состоит из n_i элементов с вероятностью исправной работы $p_{i,j}$, $j = 1, 2, \dots, n_i$, то вероятность несвязности подграфа определяется выражением:

$$q_i = 1 - \prod_{j=1}^{n_i} p_{i,j}.$$

8. Нижняя граница Эзари — Прошана. При расчете нижней границы Эзари — Прошана используется полный набор несвязных состояний сети связи (сечений).

Напомним, что G'_i , $i = 1, 2, \dots, n$ — множество всех сечений (двухполюсной сети) исходного графа G , а событие, заключающееся в том, что все элементы сечения G'_i неисправны, также обозначено как G'_i . Кроме того, пересечение обратных событий $\overline{G'_i}$ однозначно свидетельствует о связности исходного графа G :

$$p_G = P\left(\bigcap_{i=1}^n \overline{G'_i}\right).$$

Следовательно, на основе неравенства для вероятности одновременной реализации связанных случайных событий справедливо:

$$p_G \geq \prod_{i=1}^n P\left(\overline{G'_i}\right).$$

Таким образом, нижняя оценка Эзари — Прошана для вероятности связности сети имеет вид [18]:

$$p_G \geq p_e^- = \prod_{i=1}^n p_i,$$

где $p_i = P\left(\overline{G'_i}\right)$ — вероятность связности хотя бы одного элемента сечения G'_i , $i = 1, 2, \dots, n$.

Если сечение G'_i состоит из n_i элементов с вероятностью исправной работы $q_{i,j}$, $j = 1, 2, \dots, n_i$, вероятность связности хотя бы одного элемента этого сечения определяется выражением:

$$p_i = 1 - \prod_{j=1}^{n_i} q_{i,j}.$$

9. Верхняя граница Литвака — Ушакова. При расчете верхней границы Литвака — Ушакова используются только реберно-непересекающиеся сечения, то есть подмножества элементов такие, в которых какой-либо элемент не встречается дважды.

Структуры сетей связи обладают естественным свойством монотонности [19]: снижение надежности любого элемента приводит к уменьшению надежности сети. Идея данной оценки базируется на этом свойстве монотонных структур. Выделим в графе сети связи G различные наборы взаимно независимых минимальных сечений G'_i , $i = 1, 2, \dots, n$. Если в сети выделен определенный набор независимых сечений с индексами, образующими множество N_k , $k = 1, 2, \dots, N$, то для любого набора H_k вероятность связности графа, образованного этим набором, определяется как пересечение событий связности хотя бы одного элемента сечения:

$$p_k = P\left(\bigcap_{i \in N_k} \overline{G'_i}\right).$$

Для любого из наборов N_k выполняется условие $p_G \leq p_k$, то есть верхняя граница Литвака — Ушакова p_l^+ имеет вид:

$$p_G \leq p_l^+ = \min_{k=1,2,\dots,N} p_k.$$

Для нахождения вероятностей p_k учтем, что набор H_k включает m_k сечений, то есть $H_k = \{G'_{k,i}\}_{i=1,2,\dots,m_k}$, $G'_{k,i}$ — i -е сечение в k -м наборе с вероятностью связности $p_{k,i}$. Следовательно:

$$p_k = \prod_{i=1}^{m_k} p_{k,i}.$$

При этом аналогично ранее рассмотренному варианту, если сечение $G'_{k,i}$ состоит из $n_{k,i}$ элементов с вероятностью неисправной работы $q_{k,i,j}$, $j = 1, 2, \dots, n_{k,i}$, вероятность связности хотя бы одного элемента этого сечения определяется выражением:

$$p_{k,i} = 1 - \prod_{j=1}^{n_{k,i}} q_{k,i,j}.$$

В практических приложениях нет необходимости определять все возможные наборы независимых сечений. Если верхняя граница Литвака — Ушакова будет найдена не на всем множестве возможных наборов минимальных сечений, то она все же будет являться верхней границей, хотя может оказаться, что и не очень точной. Данное обстоятельство является очень важным, поскольку позволяет достаточно быстро построить неточную верхнюю границу.

10. Нижняя граница Литвака — Ушакова. При расчете нижней границы Литвака — Ушакова используются только реберно-непересекающиеся связные подграфы, то есть подмножества элементов такие, в которых какой-либо элемент не встречается дважды.

Структуры сетей связи обладают неотъемлемым свойством монотонности [19]: увеличение надежности любого элемента влечет за собой повышение надежности сети. Идея данной оценки оценок базируется на этом свойстве монотонных структур.

Выделим в графе сети связи G разные наборы взаимно независимых минимальных связных подграфов G_i , $i = 1, 2, \dots, n$, а остальные элементы удалим из сети, снизив их надежность до минимально возможного значения (нуля), и тем самым уменьшим и надежность сети в целом.

Если в сети определен некоторый набор H_k , $k = 1, 2, \dots, N$ независимых связных подграфов с индексами, образующими множество H_k , то для любого набора H_k вероятность связности графа, образованного этим набором, определяется как объединение событий связности:

$$p_k = P\left(\bigcup_{i \in N_k} G_i\right).$$

Применив правило де Моргана и определив математические ожидания обеих частей данного равенства, получаем:

$$p_k = P\left(\bigcup_{i \in N_k} G_i\right) = P\left(\overline{\bigcap_{i \in N_k} \overline{G_i}}\right) = 1 - \prod_{i \in N_k} P(\overline{G_i}).$$

Для любого из наборов N_k выполняется условие $p_G \geq p_k$, то есть нижняя граница Литвака — Ушакова p_l^- имеет вид:

$$P_G \geq P_l^- = \max_{k=1,2,\dots,N} P_k.$$

Для построения таких наборов H_k независимых связных подграфов может быть использован следующий набор рекомендаций, дающий хорошие практические результаты [19].

Так как необходимо определить наилучшую из нижних границ, то анализ начинается с кратчайших связных подграфов. Следует отметить, что для двухполюсных сетей метрикой может быть как число ребер в последовательном маршруте, так и максимальная надежность маршрута (т.е. кратчайшим путем может являться максимально надежный путь).

Таким образом, построение наборов H_k независимых связных подграфов начинается с поиска кратчайшего связного подграфа. Далее из исходного графа исключаются все элементы, содержащиеся в первом связном подграфе, и со вновь образованным графом проводится аналогичная процедура, то есть отыскивается кратчайший связный подграф, а затем его ребра удаляются. Подобный процесс выполняется до тех пор, пока граф не перестанет быть связным.

Для нахождения вероятностей p_k учтем, что набор H_k включает m_k связных подграфов, то есть $H_k = \{G_{k,i}\}_{i=1,2,\dots,m_k}$, $G_{k,i}$ — i -й связный подграф в k -м наборе с вероятностью связности $p_{k,i} = 1 - q_{k,i}$. Следовательно:

$$p_k = 1 - \prod_{i=1}^{m_k} q_{k,i}.$$

При этом аналогично ранее рассмотренному варианту, если связный подграф $G_{k,i}$ состоит из $n_{k,i}$ элементов с вероятностью исправной работы $p_{k,i,j}$, $j=1,2,\dots,n_{k,i}$, то вероятность связности этого подграфа определяется выражением:

$$p_{k,i} = \prod_{j=1}^{n_{k,i}} p_{k,i,j}.$$

Аналогично случаю наборов независимых сечений, также возможно определять негарантированную нижнюю оценку, вовлекая не все возможные наборы связных подграфов, а меньшее их число.

11. Оценки вероятностей связности магистральной сети Ростелеком «Европа — Азия». В качестве тестовой сети связи используется магистральная сеть Ростелеком, развернутая с целью формирования потоков в направлении «Европа — Азия» (рисунок 1). Отметим, что линии связи, задействованные в транзите трафика между частями света, обозначены на рисунке толстыми цветными линиями. Именно они и формируют структуру сети, которая в дальнейшем будет подвергнута анализу.

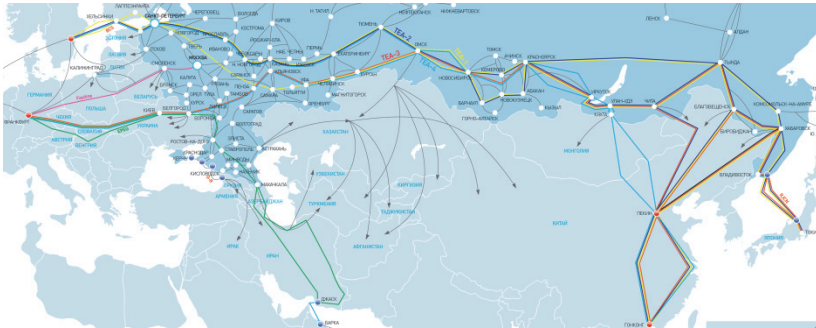


Рис. 1. Магистральная сеть Ростелеком

Помеченный граф рассматриваемой сети связи представлен на рисунке 2, а соответствие между нумерацией вершин и названиями городов приведено в таблице 1. В данной таблице полужирным шрифтом выделены вершины, являющиеся с точки зрения транспортировки трафика источниками и стоками.

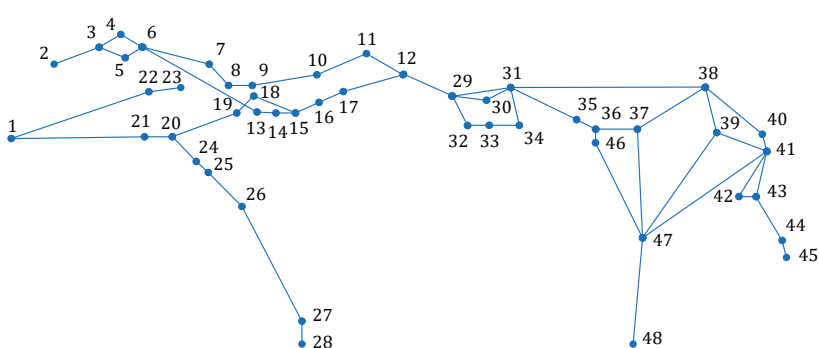


Рис. 2. Граф — модель магистральной сети Ростелеком (рисунок 1)

Таблица 1. Соответствие между нумерацией вершин и названиями городов

Номер вершины	Город	Номер вершины	Город	Номер вершины	Город
1	<i>Франкфурт</i>	17	Курган	33	Новокузнецк
2	–	18	Ульяновск	34	Абакан
3	Хельсинки	19	Пенза	35	Иркутск
4	Лаппеэнранта	20	Белгород	36	Улан-Удэ
5	–	21	Киев	37	Чита
6	Санкт-Петербург	22	Смоленск	38	Тында
7	Ярославль	23	<i>Москва</i>	39	Благовещенск
8	Нижний Новгород	24	Ростов-на-Дону	40	Комсомольск-на-Амуре
9	Казань	25	Ставрополь	41	Хабаровск
10	Екатеринбург	26	Махачкала	42	Владивосток
11	Тюмень	27	Джаск	43	–
12	Омск	28	<i>Барка</i>	44	–
13	Самара	29	Новосибирск	45	<i>Токио</i>
14	Тольятти	30	Кемерово	46	Кяхта
15	Уфа	31	Красноярск	47	Пекин
16	Челябинск	32	Барнаул	48	<i>Гонконг</i>

Вероятность связности p_{ij} для каждой линии связи задается как ее коэффициент готовности и вычисляется для волоконно-оптических систем передачи на основе протяженности d [20]:

$$p_{i,j} = 0,99987^{\frac{d}{402,325}}, \quad i, j = 1, 2, \dots, 48.$$

Протяженность d определяется на основе географических координат городов, между которыми развернута данная линия связи. Рассчитанные параметры линий связи приведены в таблице 2.

В качестве примера на рисунке 3 представлены зависимости вероятности связности p_G графа от количества i учтенных подграфов для направления связи 2–48 в случае использования простых цепей и сечений.

Для этого же направления связи 2–48 на рисунке 4 показаны зависимости граничных оценок вероятности связности p_G графа от количества i учтенных подграфов. Отметим, что нижняя граница Эзари — Прошана и верхняя Литвака — Ушакова практически совпали, что делает возможным практически близкую к точной оценку вероятности связности на основе данных границ.

Таблица 2. Параметры линий связи

Линия $i - j$	Протяженность d , км	Вероятность связности p_{ij}	Линия $i - j$	Протяженность d , км	Вероятность связности p_{ij}	Линия $i - j$	Протяженность d , км	Вероятность связности p_{ij}
2 – 3	630	0,9998	19 – 20	945	0,99969	35 – 36	135	0,99996
3 – 5	360	0,99988	20 – 21	405	0,99987	36 – 37	585	0,99981
3 – 4	315	0,9999	1 – 21	1710	0,99945	37 – 38	1035	0,99967
4 – 6	315	0,9999	1 – 22	1890	0,99939	31 – 38	2295	0,99926
5 – 6	180	0,99994	22 – 23	450	0,99985	38 – 39	630	0,9998
6 – 7	945	0,99969	20 – 24	495	0,99984	39 – 41	675	0,99978
7 – 8	360	0,99988	24 – 25	180	0,99994	38 – 40	990	0,99968
8 – 9	315	0,9999	25 – 26	675	0,99978	40 – 41	225	0,99993
9 – 10	900	0,99971	26 – 27	1710	0,99945	41 – 42	675	0,99978
10 – 11	675	0,99978	27 – 28	315	0,9999	42 – 43	135	0,99996
11 – 12	585	0,99981	12 – 29	720	0,99977	41 – 43	630	0,9998
12 – 17	855	0,99972	29 – 30	450	0,99985	43 – 44	720	0,99977
16 – 17	315	0,9999	30 – 31	360	0,99988	44 – 45	180	0,99994
15 – 16	360	0,99988	29 – 31	765	0,99975	36 – 46	135	0,99996
14 – 15	180	0,99994	29 – 32	450	0,99985	46 – 47	1395	0,99955
13 – 14	270	0,99991	32 – 33	270	0,99991	37 – 47	1350	0,99956
6 – 13	1755	0,99943	33 – 34	270	0,99991	39 – 47	1710	0,99945
15 – 18	495	0,99984	31 – 34	405	0,99987	41 – 47	1980	0,99936
18 – 19	315	0,9999	31 – 35	945	0,99969	47 – 48	1440	0,99953

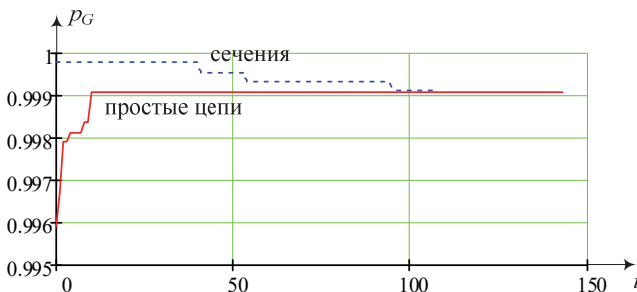


Рис. 3. Зависимости вероятности связности p_G графа от количества i учтенных подграфов для направления связи 2–48 (рисунки 1, 2)

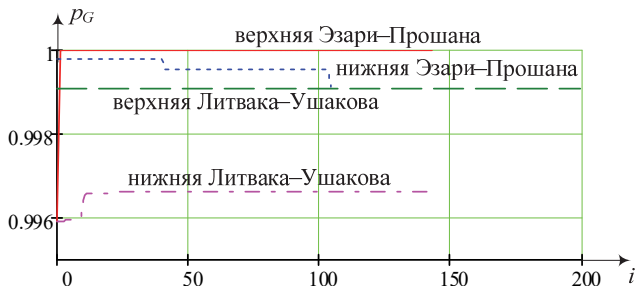


Рис. 4. Зависимости граничных оценок вероятности связности p_G графа от количества i учтенных подграфов для направления связи 2–48 (рисунки 1, 2)

В таблице 3 приведены сводные данные по точным и граничным оценкам вероятности связи для различных направлений и для графа в целом, а также информация о количестве анализируемых подграфов (простых цепей, деревьев, сечений) и временных затрат на расчет соответствующих параметров. Отметим, что реализация вышеописанных процедур анализа выполнена в программной среде MathCad 15, а вычисления проводились на компьютере с процессором Intel(R) Core(TM) i5-3470 CPU @ 3,20 ГГц 3,60 ГГц, оперативной памятью 8 ГБ и 64-разрядной операционной системой Windows 7. Отметим, что применение метода полного перебора оказалось малоэффективным для данной сети (рисунок 1), поскольку потребовало проверки 2^{57} состояний.

Таблица 3. Результирующие данные анализа устойчивости магистральной сети Ростелеком

Направление связи	Тип оценки	Число подграфов	Вероятность связности	Время анализа, с
1–23 (Франкфурт – Москва)	верхняя граница Эзари – Прошана	1	0,999 244	0,06
	верхняя граница Литвака – Ушакова	2	0,999 244	0,109
	точная на основе путей	1	0,999 244	0,09
	точная на основе сечений	2	0,999 244	0,09
	нижняя граница Эзари – Прошана	2	0,999 244	0,08
1–28 (Франкфурт – Барка)	нижняя граница Литвака – Ушакова	1	0,999 244	0,06
	верхняя граница Эзари – Прошана	1	0,998 228	0,05
	верхняя граница Литвака – Ушакова	9	0,998 228	0,28
	точная на основе путей	1	0,998 228	0,09
	точная на основе сечений	9	0,998 228	0,28
	нижняя граница Эзари – Прошана	9	0,997 966	0,28
нижняя граница Литвака – Ушакова	1	0,998 228	0,05	

Продолжение таблицы 3.

Направление связи	Тип оценки	Число подграфов	Вероятность связности	Время анализа, с
1-45 (Франкфурт – Токио)	верхняя граница Эзари – Прошана	216	1	0,06
	верхняя граница Литвака – Ушакова	153	0,998 255	6,66
	точная на основе путей	216	0,998 226	5,55
	точная на основе сечений	153	0,998 226	52,54
	нижняя граница Эзари – Прошана	153	0,996 891	1,89
	нижняя граница Литвака – Ушакова	216	0,996 154	0,08
1-48 (Франкфурт – Гонконг)	верхняя граница Эзари – Прошана	72	1	0,05
	верхняя граница Литвака – Ушакова	158	0,998 98	7,11
	точная на основе путей	72	0,998 502	0,52
	точная на основе сечений	158	0,998 502	43,66
	нижняя граница Эзари – Прошана	158	0,996 717	1,05
	нижняя граница Литвака – Ушакова	72	0,996 473	0,05
2-23 («←» – Москва)	верхняя граница Эзари – Прошана	4	1	0,11
	верхняя граница Литвака – Ушакова	76	0,998 69	5,85
	точная на основе путей	4	0,997 791	0,13
	точная на основе сечений	76	0,997 791	1,48
	нижняя граница Эзари – Прошана	76	0,996 398	1,44
	нижняя граница Литвака – Ушакова	4	0,996 907	0,23
2-28 («←» – Барка)	верхняя граница Эзари – Прошана	4	1	0,23
	верхняя граница Литвака – Ушакова	40	0,998 14	1,06
	точная на основе путей	4	0,998 139	0,12
	точная на основе сечений	40	0,998 139	0,75
	нижняя граница Эзари – Прошана	40	0,998 139	0,87
	нижняя граница Литвака – Ушакова	4	0,997 255	0,23
2-45 («←» – Токио)	верхняя граница Эзари – Прошана	432	1	0,18
	верхняя граница Литвака – Ушакова	100	0,999 273	6,57
	точная на основе путей	432	0,999 271	32,33
	точная на основе сечений	100	0,999 271	124,2
	нижняя граница Эзари – Прошана	100	0,999 271	1,65
	нижняя граница Литвака – Ушакова	432	0,996 313	0,3

Продолжение таблицы 3.

Направление связи	Тип оценки	Число подграфов	Вероятность связности	Время анализа, с
2–48 («–» – Гонконг)	верхняя граница Эзари – Прошана	144	1	0,14
	верхняя граница Литвака – Ушакова	105	0,999 098	6,71
	точная на основе путей	144	0,999 097	2,8
	точная на основе сечений	105	0,999 097	201,6
	нижняя граница Эзари – Прошана	105	0,999 097	0,87
	нижняя граница Литвака – Ушакова	144	0,996 632	0,14
23–28 (Моск-ва – Барка)	верхняя граница Эзари – Прошана	1	0,997 473	0,05
	верхняя граница Литвака – Ушакова	15	0,997 473	0,37
	точная на основе путей	1	0,997 473	0,09
	точная на основе сечений	15	0,997 473	0,5
	нижняя граница Эзари – Прошана	1	0,995 358	0,34
	нижняя граница Литвака – Ушакова	15	0,997 473	0,05
23–45 (Моск-ва – Токио)	верхняя граница Эзари – Прошана	216	1	0,2
	верхняя граница Литвака – Ушакова	159	0,997 5	7,02
	точная на основе путей	216	0,997 472	5,71
	точная на основе сечений	159	0,997 472	54,26
	нижняя граница Эзари – Прошана	159	0,993 591	2,03
	нижняя граница Литвака – Ушакова	216	0,995 401	0,065
23–48 (Моск-ва – Гонконг)	верхняя граница Эзари – Прошана	72	1	0,05
	верхняя граница Литвака – Ушакова	164	0,997 326	7,52
	точная на основе путей	72	0,997 298	0,64
	точная на основе сечений	164	0,997 298	51
	нижняя граница Эзари – Прошана	164	0,993 418	1,2
	нижняя граница Литвака – Ушакова	72	0,995 719	0,05
28–45 (Барка – Токио)	верхняя граница Эзари – Прошана	216	1	0,06
	верхняя граница Литвака – Ушакова	129	0,998 053	6,68
	точная на основе путей	216	0,997 82	6,01
	точная на основе сечений	129	0,997 82	41,25
	нижняя граница Эзари – Прошана	129	0,996 369	2,09
	нижняя граница Литвака – Ушакова	216	0,995 748	0,08

Продолжение таблицы 3.

Направле- ние связи	Тип оценки	Число подграфов	Вероят- ность связности	Время анализа, с
28–48 (Барка – Гонконг)	верхняя граница Эзари – Прошана	72	1	0,05
	верхняя граница Литвака – Ушакова	134	0,997 792	7,15
	точная на основе путей	72	0,997 646	0,51
	точная на основе сечений	134	0,997 646	35,83
	нижняя граница Эзари – Прошана	134	0,996 196	1,19
	нижняя граница Литвака – Ушакова	72	0,996 067	0,05
45–48 (Токио – Гонконг)	верхняя граница Эзари – Прошана	22	1	0,37
	верхняя граница Литвака – Ушакова	33	0,999 794	24,16
	точная на основе путей	22	0,999 244	0,44
	точная на основе сечений	33	0,999 244	25,13
	нижняя граница Эзари – Прошана	33	0,998 736	24,68
	нижняя граница Литвака – Ушакова	22	0,998 402	0,39
Сеть в целом	верхняя граница Эзари – Прошана	1 096 704	–	–
	верхняя граница Литвака – Ушакова	190	0,995 718	10,37
	точная на основе путей	1 096 704	–	–
	точная на основе сечений	190	–	–
	нижняя граница Эзари – Прошана	190	0,995 716	2,53
	нижняя граница Литвака – Ушакова	1 096 704	–	–

Данная таблица наглядно демонстрирует преимущество использования точных оценок на основе связанных состояний, не смотря на их большее количество по сравнению с несвязными состояниями, что является следствием особенностей анализируемого графа, которые проявляются в возможности его представления в виде совокупности подграфов и имеют лишь одну точку сочленения. Результатом этого же оказывается вычислительная сложность расчета верхних границ Литвака — Ушакова, поскольку требует рассмотрения огромного количества (более 10^5) независимых наборов сечений.

Расчет точных значений вероятности связности для графа в целом оказался практически невозможным в связи с нехваткой оперативной памяти. В то же время вычисления на основе тысячи деревьев заняли около 105 секунд и привели к величине вероятности связности $p_G = 0,992\ 553$, а на основе ста шестнадцати сечений — 27 секунд и

$p_G = 0,996\ 469$. Верхняя граница Эзари — Прошана и нижняя Литвака — Ушакова также оказались недоступными для расчетов в виду значительного числа остовых деревьев (более одного миллиона), однако границы, использующие сечения показали величины вероятности связности, различные лишь в шестом знаке после запятой.

12. Заключение. В заключение следует отметить, что, несмотря на существенный недостаток метода полного перебора типовых состояний, заключающийся в значительной трудоемкости проводимых вычислений, он оказывается достаточно востребованным, поскольку наглядно и естественно описывает понятия связности и несвязности сети, что приводит к существенному упрощению реализации вычислительных процедур, используемых в рамках анализа устойчивости. Преимущества метода перебора типовых состояний по сравнению с полным перебором всех состояний позволяют применять рассмотренный метод на более сложноразветвленных сетях, характеризующихся значительными мощностями как вершин, так и ребер в графе. Кроме того, на основе метода полного перебора типовых состояний оказывается возможным получать граничные оценки вероятности связности сети, причем одни из них, а именно границы Литвака — Ушакова, позволяют ограничиваться перебором неполного числа вариантов наборов независимых связных и несвязных состояний графа, что также существенно снижает вычислительные сложности решаемых задач.

Следует подчеркнуть, что точные оценки устойчивости сетей, полученные на основе рассмотренного метода, обладают, с одной стороны, приемлемой вычислительной сложностью, а с другой — достаточной достоверностью. Так, практически для всех рассмотренных вариантов направлений связи (таблица 3) точные оценки в сравнении с приближенными имеют различия в третьем знаке после запятой, что является критическим, поскольку требования к устойчивости (надежности и живучести), предъявляемые существующим ГОСТ [7], разнятся во втором и третьем знаке.

Литература

1. Zuev K.M., Wu S., Beck J.L. Network reliability problem and its efficient solution by Subset Simulation // Probabilistic Engineering Mechanics. vol. 40. pp. 25–35.
2. Mussel C., Hopfensitz M., Kestler H.A. Boolnet package vignette // 2015. 49 p.
3. Teruggia R. Reliability Analysis of Probabilistic Networks: PhD Thesis // University of Turin. 2010. 241 p.
4. Дудник Б.Я., Овчаренко В. Ф. Надежность и живучесть систем связи // М.: Радио и связь. 1984. 216 с.
5. Филин Б.П. Методы анализа структурной надежности сетей связи // Москва : Радио и связь. 1988. 208 с.

6. *Половко А.М., Гуров С.В.* Основы теории надежности // БХВ-Петербург. 2006. 704 с.
7. ГОСТ Р 53111–2008. Устойчивость функционирования сети связи общего пользования. Требования и методы проверки. // М. : Стандартинформ. 2009. 16 с.
8. *Обоскалов В.П.* Структурная надежность электроэнергетических систем: учеб. пособие // Екатеринбург: УрФУ. 2012. 194 с.
9. *Батенков К.А.* Числовые характеристики структур сетей связи // Труды СПИИРАН. 2017. № 4(53). С. 5–28.
10. *Chaturvedi S.K.* Network Reliability, Measures and Evaluation // John Wiley & Sons. 2016. 237 p.
11. *Tamm V.* Теория графов // М.: Мир. 1988. 424 с.
12. *Zhao J., Yagan O., Gligor V.* On the Strengths of Connectivity and Robustness in General Random Intersection Graphs // 53rd IEEE Conference on Decision and Control. 2014. pp. 3661–3668.
13. *Yagan O., Makowski A.M.* Zero-one Laws for Connectivity in Random Key Graphs // IEEE Transactions on Information Theory. 2012. vol. 58. no. 5. pp. 2983–2999.
14. *Zhao J., Yagan O., Gligor V.* Connectivity in Secure Wireless Sensor Networks under Transmission Constraints // 2014 52nd Annual Allerton Conference on Communication, Control, and Computing. 2014. pp. 1294–1301.
15. *Nuñez A. et al.* Detecting series periodicity with horizontal visibility graphs // International Journal of Bifurcation and Chaos. 2012. vol. 22. no. 07. pp. 1250160.
16. *Zhang H.C. et al.* Connection effect on amplitude death stability of multi-module floating airport // Ocean Engineering. 2017. vol. 129. pp. 46–56.
17. *Батенков К.А.* Устойчивость сетей связи // Академия ФСО России. 2017. 277 с.
18. *Brown J.I., Tufts J.* On the roots of domination polynomials // Graphs and Combinatorics. 2014. vol. 30. no. 3. pp. 527–547.
19. *Ушаков И.А.* Курс теории надежности систем // М.: Дрофа. 2008. 239 с.
20. *Pino W., Gomes T., Kooij R.* A Comparison between Two All-Terminal Reliability Algorithms // Journal of Advances in Computer Networks. 2015. vol. 3. no. 4. pp. 284–290.

Батенков Кирилл Александрович — д-р техн. наук, доцент, сотрудник, Академия Федеральной службы охраны Российской Федерации. Область научных интересов: статистическая теория связи, модели и методы обработки сигналов, анализ качества сетей связи. Число научных публикаций — 150. pustur@yandex.ru; ул. Приборостроительная, 35, 302015, Орёл, Российская Федерация; р.т.: +7(486)254-13-25.

K.A. BATENKOV
**ACCURATE AND BOUNDARY ESTIMATE OF
COMMUNICATION NETWORK CONNECTIVITY PROBABILITY
BASED ON MODEL STATE COMPLETE ENUMERATION
METHOD**

Batenkov K.A. Accurate and Boundary Estimate of Communication Network Connectivity Probability Based on Model State Complete Enumeration Method.

Abstract. We consider one of communication network structure analysis and synthesis methods, based on the simplest approach to connectivity probability calculation – a method of full network typical state search. In this case, the typical states of the network are understood as the events of network graph connectivity and disconnection, which are simple graph chains and sections. Despite significant drawback of typical state enumeration method, which involves significant calculation complexity, it is quite popular at stage of debugging new analysis methods. In addition, on its basis it is possible to obtain boundary estimates of network connectivity probability. Thus, when calculating Asari–Proshana boundaries use full set of incoherent (top) and cohesive (bottom) communication network states. These boundaries are based on statement that network connectivity probability under same conditions is higher (lower) than that of network composed of independent disjoint (connected) subgraph complete set serial (parallel) connection. When calculating Litvak–Ushakov boundaries, only edge-disjoint sections (for upper) and connected subgraphs (for lower) are used, i.e. subsets of elements such that any element does not meet two-rods. This boundary takes into account the well-known natural monotonicity property, which is to reduce (increase) network reliability with decrease (increase) any element reliability. From a computational view point Asari–Proshana boundaries have huge drawback: they require references of all connected subgraphs to compute upper bounds and all minimal cuts for bottom, which in itself is non-trivial. Litvak–Ushakov boundaries are devoid of these drawback: by calculating them, we can stop at any searching step for variants of sets of independent connected and disconnected graph states.

Keywords: Network, Graph Structure, Probability of Connectedness, Model State Complete Enumeration Method, Boundary Evaluation.

Batenkov Kirill Aleksandrovich — Ph.D., Dr.Sci., Associate Professor, Researcher, Academy of Federal Security Guard Service of the Russian Federation. Research interests: statistical communication theory, models and methods of signal processing, communications network quality analysis. The number of publications — 150. pustur@yandex.ru; 35, Priborostroitel'na-ya str., 302015, Orel, Russian Federation; office phone: +7(486)254-13-25.

References

1. Zuev K.M., Wu S., Beck J.L. Network reliability problem and its efficient solution by Subset Simulation. *Probabilistic Engineering Mechanics*. vol. 40. pp. 25–35.
2. Mussel C., Hopfensitz M., Kestler H.A. Boolnet package vignette. 2015. 49 p.
3. Teruggia R. Reliability Analysis of Probabilistic Networks: PhD Thesis. University of Turin. 2010. 241 p.
4. Dudnik B. Ya., Ovcharenko V. F. *Nadezhnost' i zhivuchest' sistem svyazi* [The reliability and survivability of communication systems]. M.: Radio i svyaz'. 1984. 216 p. (In Russ.).
5. Filin B P. *Metody analiza strukturnoj nadezhnosti setej svyazi* [Methods of analysis of structural reliability of communication networks]. M.: Radio i svyaz'. 1988. 208 p. (In Russ.).

6. Polovko A.M., Gurov S.V. *Osnovy teorii nadezhnosti* [Fundamentals of reliability theory]. BHV-Peterburg. 2006. 704 p. (In Russ.).
7. GOST R 53111–2008. *Ustojchivost' funkcionirovaniya seti svyazi obshchego pol'zovaniya. Trebovaniya i metody proverki* [Stability of the public communication network. Requirements and verification methods]. M.: Standartinform. 2009. 16 p. (In Russ.).
8. Oboskalov V.P. *Structural reliability of electric power systems* [Structural reliability of electric power systems]. Ekaterinburg: UrFU. 2012. 194 p. (In Russ.).
9. Batenkov K.A. [Numerical characteristics of the structures of communication networks]. *Trudy SPIIRAN – Proceedings of SPIIRAS*. 2017. vol. 4(53). pp. 5–28. (In Russ.).
10. Chaturvedi S.K. Network Reliability. Measures and Evaluation. John Wiley & Sons. 2016. 237 p.
11. Tutte W.T. Graph Theory. AddisonWesley Publishing Company. 1984. 333 p. (Russ. ed.: Tatt U. *Teoriya grafov*. M.: Mir. 1988. 424 p.).
12. Zhao J., Yagan O., Gligor V. On the Strengths of Connectivity and Robustness in General Random Intersection Graphs. 53rd IEEE Conference on Decision and Control. 2014. pp. 3661–3668.
13. Yagan O. Makowski A.M. Zero-one Laws for Connectivity in Random Key Graphs. *IEEE Transactions on Information Theory*. 2012. vol. 58. no. 5. pp. 2983–2999.
14. Zhao J., Yagan O., Gligor V. Connectivity in Secure Wireless Sensor Networks under Transmission Constraints. 2014 52nd Annual Allerton Conference on Communication, Control, and Computing. 2014. pp. 1294–1301.
15. Nuñez A. et al. Detecting series periodicity with horizontal visibility graphs. *International Journal of Bifurcation and Chaos*. 2012. vol. 22. no. 07. pp. 1250160.
16. Zhang H.C. et al. Connection effect on amplitude death stability of multi-module floating airport. *Ocean Engineering*. 2017. vol. 129. pp. 46–56.
17. Batenkov K. A. *Ustojchivost' setej svyazi* [Network stability]. Akademiya FSO Rossii. 2017. 277 p. (In Russ.).
18. Brown J.I., Tufts J. On the roots of domination polynomials. *Graphs and Combinatorics*. 2014. vol. 30. no. 3. pp. 527–547.
19. Ushakov I.A. *Kurs teorii nadezhnosti sistem* [Course in the theory of reliability systems]. M.: Drofa. 2008. 239 p. (In Russ.).
20. Pino W., Gomes T., Kooij R. A Comparison between Two All-Terminal Reliability Algorithms. *Journal of Advances in Computer Networks*. 2015. vol. 3. no. 4. pp. 284–290.

Ю.В. Доронина, А.В. Скатков
**АНАЛИЗ СТАТИСТИЧЕСКОЙ УСТОЙЧИВОСТИ
СТАЦИОНАРНЫХ МАРКОВСКИХ МОДЕЛЕЙ**

Доронина Ю.В., Скатков А.В. Анализ статистической устойчивости стационарных Марковских моделей.

Аннотация. Предложен подход для оценки качества стационарных Марковских моделей без поглощающих состояний на основе меры статистической устойчивости: формулируется описание меры и определяются ее свойства. Показано, что оценки статистической устойчивости моделей описывались разными авторами либо как методологический аспект качества модели, либо в рамках других модельных свойств. При решении практических задач имитационного моделирования, например на основе Марковских моделей, возникает выраженная проблема обеспечения размерности требуемых выборок. На основе введенных формулировок предложен конструктивный подход к решению задач оптимизации объема выборки и анализа статистической волатильности Марковской модели к возникающим аномалиям при ограничениях на точность результатов, что обеспечивает требуемую достоверность и исключение нефункциональной избыточности.

Для анализа вида переходов в матрице переходов введена мера ее дивергенции (нормированная и центрированная). Эта мера не обладает полнотой описания и используется в качестве иллюстративной характеристики моделей определенного свойства. Оценка дивергенции матриц переходов может быть полезна при исследовании моделей с высокой чувствительностью обнаружения исследуемых свойств объектов. Сформулированы ключевые этапы подхода, который связан с исследованием квазиоднородных моделей.

На примере моделирования реального технического объекта с отказами, восстановлениями и профилактикой предложены количественные оценки статистической устойчивости и статистической волатильности модели. Показана эффективность предлагаемых подходов при решении задачи анализа статистической устойчивости в задачах квалитетного анализа квазиоднородной модели сложных систем. На основе предложенного конструктивного подхода получен оперативный инструмент принятия решений по параметрической и функциональной настройке сложных технических объектов на долгосрочную и краткосрочную перспективы.

Ключевые слова: статистическая устойчивость, квазиоднородная модель, статистическая волатильность, случайное блуждание, цепь Маркова, сложная техническая система, точность, вероятность перехода, число реализаций процесса, дивергенция модели.

1. Введение. Исследование процесса функционирования сложных системотехнических комплексов в большинстве случаев сталкивается с рядом объективных проблем: неопределенность состояний, размерность моделей, отсутствие возможности натурных испытаний, высокая стоимость получения данных, невозможность мониторинга состояния вследствие сложных условий функционирования системы и тому подобное. В этих случаях модельно-ориентированный подход в рамках имитационного моделирования оказывается единственно возможным средством исследования системы [1, 2].

Область задач имитационного моделирования включает четыре основных направления: моделирование динамических систем, дискретно-событийное моделирование, моделирование системной

динамики, агентное моделирование. В каждом из этих направлений может быть применена Марковская модель (ММ) (в агентном моделировании косвенно) в форме представления и анализа Марковских событий, Марковской последовательности или Марковского случайного процесса (МСП) [3, 4].

Анализ процесса функционирования СТО должен быть соотнесен с этапами жизненного цикла системы: на начальных этапах и на этапе деградации оценки свойств СТО могут значительно отличаться. Применение ММ к анализу надежности и прогнозированию временных характеристик СТО широко распространено [5-7], но не во всех случаях результаты оказываются достоверными, в связи с этим целесообразно развивать механизмы решения задач анализа функционирования СТО с учетом оценки качества решений в рамках ММ (таблица 1).

Таблица 1. Некоторые задачи и модели исследования СТО

Задача	Учет статистической устойчивости, метрики, ссылки
Из теории надежности: для нормального функционирования изучаемой системы нужно, чтобы параметры, характеризующие систему во время ее работы, не выходили за некоторые пределы.	Исследуются границы выхода Марковского процесса за заданные границы ($y+dy$) с точностью $o(dy)$ [8].
Закон распределения времени пребывания МСП в заданной области.	1. Если за начальный момент времени взят момент пересечения значениями случайного процесса (СП) границы допустимой области, то устанавливается закон распределения времени пребывания значений этого СП в допустимой области. 2. Устанавливается закон распределения времени выброса значений рассматриваемого СП за уровень «снизу вверх» [8].
Определение среднего числа выбросов значений МСП в единицу времени (время пребывания вне допустимой области).	Сводится к решению соответствующих задач в рамках уравнений Колмогорова.
Анализ надежности систем с определением вероятностей пребывания в состояниях для различных режимов работы системы: стационарный, нестационарный.	Определяются финальные вероятности, находится распределение вероятностей до стационарного режима [9, 10].
Задачи оценки параметров восстановления и обслуживания систем.	Нахождение параметров эффективности СМО; распределение времен пребывания в состояниях, показатели надежности систем [11].

Известно, что вопросам оценки результатов имитационного моделирования уделялось много внимания, использовались критерии Стьюдента, Фишера, Манна — Уитни, Вилкоксона и другие. Но в существующих исследованиях недостаточно проработаны вопросы устойчивости моделей. В то же время устойчивость — это фундаментальное свойство динамических систем, которое исследуется обычно в двух плоскостях: как реакция системы на внешние возмущения динамического характера, так и изменение параметров в ответ на эти возмущения. В представленном исследовании предложен некоторый подход к оценке статистической устойчивости (СУ) ММ в задачах анализа системной динамики, в котором не детализируется источник возникновения неустойчивости, а анализируются свойства модели в ее выходных характеристиках. Предложенный подход отличается конструктивизмом, позволяя исследователю формулировать некоторые оценки СУ, в том числе на основе внешней монотонности ММ и модельной волатильности.

2. Обзор исследований статистической устойчивости моделей. Основные результаты связаны с анализом качества моделей в рамках квалиметрии моделей. Как отдельный показатель качества модели, статистическая устойчивость позволяет исследовать статистики (функций выборки), в том числе стабильности частоты (относительной частоты) массовых событий и средних величин (т.е. слабой их зависимости от объема данных) [13, 14]. В работе [14] предложено выполнять сопоставление модели объекту по целостному свойству через измерение расстояния $\rho(M, ob)$ между соответствующими точками в n -мерном пространстве элементарных свойств объекта. В качестве аргумента в соответствующих функционалах рассматривают величину разности выходных состояний объекта $y_j(Ob)$ и модели $y_j(M)$. В этом случае условие адекватности $|y_j(Ob) - y_j(M)| \leq \varepsilon$ по n свойствам обобщается в условие $\rho(M, ob) \leq \varepsilon, \varepsilon > 0$. Для измерения расстояния $\rho(M, ob)$ могут использоваться различные метрики. Эти величины рекомендовано использовать для оценки моделей по адекватности.

В монографии [12] отражено, что не всегда случайные модели в полной мере отражают специфику реальных событий, и на больших интервалах наблюдения спекавали, что гипотеза идеальной статистической устойчивости не находит экспериментального подтверждения. На больших объемах данных уровень флуктуаций реальных статистик практически не изменяется или возрастает. Это связано с тем, что не удастся обеспечить повторяемость условий экспериментов.

Исследования авторов [15, 16] связаны с процедурой многокритериального структурно-функционального синтеза комплекса разнотипных моделей, описывающих с различной степенью детализации разные аспекты функционирования системы проактивного управления группировкой СТО в динамически изменяющейся обстановке, задаваемой как стохастическими и интервальными исходными данными, так и данными, имеющими нечетко-возможностный характер; позволившая на конструктивном уровне количественно оценить робастность и устойчивость программ проактивного управления СТО на основе построения и аппроксимации областей достижимости логико-динамических моделей, которые описывают структурную динамику рассматриваемых СТО.

Таким образом, вопросы оценки статистической устойчивости моделей поднимались разными авторами либо как методологический аспект качества модели, либо в рамках других модельных свойств. Однако в практических задачах имитационного моделирования, например на основе ММ, возникает выраженная проблема размерности выборки.

Всюду далее рассматривается ММ без поглощающих состояний в рамках стационарного (квазистационарного) режимов, S_0 — стартовое состояние системы; S_n — n -ое состояние, в котором оказывается система после n -ого перехода; символом " $\xrightarrow{M_x}$ " будем обозначать отображение, осуществляемое имитационной моделью M_x ; M_x, M_v, M_m — обозначения имитационных моделей соответственно: общее обозначение имитационной модели, модель с результатом в виде вектора, модель с результатом в виде матрицы.

Обобщая опыт, который был накоплен при ММ сложных систем, и учитывая предложенную в [17] классификацию математических схем моделирования, выделим три основных способа построения таких моделей (видов математических схем моделирования):

$$schema P_v : \left\| P_{ij} \right\| \xrightarrow{M_v} \langle S_0, S_1, \dots, S_n \rangle; \quad (1)$$

$$schema P_m : \langle S_0, S_1, \dots, S_n \rangle \xrightarrow{M_v} \left\| P_{ij} \right\| \xrightarrow{M_m} \left\| P_{ij}^k \right\|; \quad (2)$$

$$schema A : schema P_v \xrightarrow{M_x} schema P_m; \quad (3)$$

$$schema A : \left\| P_{ij} \right\| \xrightarrow{M_v} \left\| P_{ij}^{M_x} \right\| \xrightarrow{M_m} \left\| \theta \right\|,$$

где P_v — дискретно-стохастическая модель (P -схема), модель с ожидаемым результатом в виде вектора; P_m — дискретно-стохастическая модель (P -схема), модель с ожидаемым результатом в виде матрицы; A — комбинированная модель (A -схема); $\|P_{ij}\|$ — матрица переходов ММ; $\langle S_0, S_1, \dots, S_n \rangle$ — последовательность состояний ММ; k — число запусков модели, за которое система из первоначального состояния перейдет в терминальное состояние, описываемое матрицей $\|P_{ij}^k\|$; $\|P_{ij}^{Mm}\|$ — «восстановленная» матрица переходных вероятностей, содержащая статистические оценки P_{ij} , (именно эта матрица будет использована для анализа статистической устойчивости модели M_x); $\|\theta\|$ — матрица результатов моделирования, (например: финальные вероятности переходов; некоторая мера отклонения числа совпавших элементов $\|P_{ij}\|$ и $\|P_{ij}^{Mm}\|$ и т.п.).

Схему P_v рекомендуется использовать в случаях, когда необходимо изучить динамику системных состояний в форме марковских переходов $\langle S_0, S_{i-1}, S_i, S_{i+1} \dots \rangle$, где S_i — одно из состояний, в которое система может перейти из состояния S_{i-1} . Схема P_m позволяет на основе известной динамики системы получить оценку случайных величин — вероятностей переходов. Схема С реализуется на основе агрегации схемы P_v и схемы P_m и ее удобно использовать для исследования процессов сходимости процессов моделирования.

Вопросы планирования экспериментов исследовались многими авторами, один из возможных подходов будем использовать в следующем виде. Планы имитационных экспериментов по приведенным схемам P_v , P_m , A (1)-(3) ММ могут быть представлены в виде кортежа $\Pi \langle n, l, k, \theta \rangle$, где n — скаляр, задающий число сгенерированных цепочек ММ в выражении (1), $n = |N|$; l_m — длина m -ой цепи Маркова, $l = |L|$; k — кратность запусков модели, $k = |K|$; $\theta_{n,l,k}$ — многомерная матрица, которая может содержать результаты первичной статистики, на основе которой будет оценена статистическая устойчивость. Таким образом, если обозначить D — множество допустимых планов ($J = |D|$), то $\Pi_j \langle n_j, l_j, k_j, \theta_j \rangle$ — реализуемый план.

Использование формализации (1)-(3) позволит алгоритмизировать процесс анализа статистической устойчивости квазиоднородной стационарной ММ без поглощающих состояний. Введем некоторые определения и нотации на основе имеющегося опыта и существующих исследований.

3. Основные определения. Под качеством модели будем понимать совокупность свойств и характеристик модели, определяющих его способности удовлетворять обусловленные или предполагаемые потребности субъекта, использующего модель [18, 19]. В различных источниках различают несколько аспектов понятия «устойчивость». Устойчивость модели по отношению к изменениям ее параметров означает сохранение аппарата моделирования, основных связей между переменными, типов ограничений в некотором интервале ее параметров [20, 21].

Другим аспектом устойчивости является устойчивость решения задачи (результатов) моделирования (обнаруженных свойств, сценариев, траекторий, состояний) по отношению к изменениям параметров модели или начальных условий. Если зависимость от параметров и начальных условий является регулярной, то малые ошибки в исходных данных приведут к небольшому изменению результата. Тогда, решая, например, задачу выбора по приближенному данным, можно обоснованно говорить о нахождении приближенного решения. Иногда устойчивость является целью практического моделирования. В частности — поиск алгоритмов деятельности человека без разрушения природной экосистемы [10]. Практически устойчивость статистик трактуется исследователем самым различным образом, как правило, используются оценки повторяемости и монотонное уменьшение длины доверительного интервала по мере роста числа реализаций при фиксированной доверительной вероятности.

Планы экспериментов Π по схемам P_v, P_m, A (1)-(3) порождают вариативность процессов исследования моделей, и, как следствие, значительный разброс статистических оценок, особенно при решении задач со сложной системной динамикой.

Мера *статистической устойчивости (СУ)* модели M_x может быть определена различными способами в зависимости от целей исследования и типа исследуемого объекта. В качестве оценочного параметра примем (как оценку текущей статистики модели M_x), представляющего интерес для исследования статистики по полученным результатам экспериментов в виде вектора θ . Пусть $\tilde{\gamma}$ — заданная ЛППР оценка статистики модели M_x ; Δ_γ — мера близости текущей и заданной оценок статистики модели M_x ; k — число запусков модели; ε — точность модели; p — доверительная вероятность оценки параметров.

Определение 1. Модель M_x , реализованная в рамках одной из видов математических схем моделирования P_v, P_m, A (1)–(3) является статистически устойчивой, если для квалитметрической оценки ее устойчивости, начиная с некоторых $n_i \geq n_0, l_i \geq l_0, k_i \geq k_0$, будет выполняться условие:

$$\left| \gamma^{M_x}(\theta(n_i, l_i, k_i)) - \tilde{\gamma} \right|_{\varepsilon, p} \leq \Delta\gamma \quad (4)$$

с учетом условия сходимости по статистической вероятности.

Следовательно, модель M_x может быть определена в общем случае как k — статистически устойчивая ($n-, l-, \varepsilon-, p$ — статистически устойчивая) на основе заданного плана экспериментов $\Pi(n, l, k, \theta)$ согласно определению 1, (4).

Таким образом, если при исследовании СТО на основе ММ по схеме: *schema A*: $\|P_{ij}\| \xrightarrow{M_v} \|P_{ij}^{M_x}\| \xrightarrow{M_m} \|\theta\|$ отклонение $\|P_{ij}^{M_m}\|$ в модели M_m не превысит заданную ЛПР величину $\tilde{\theta}$, то для плана экспериментов $\Pi(n, l, k, \theta)$ модель M_x будет считаться $\theta(k-, \varepsilon-, p-)$ — статистически устойчивой и обозначаться $\gamma_k^{M_x}, \gamma_\varepsilon^{M_x}, \gamma_p^{M_x}$.

Формализуем основные свойства меры статистической устойчивости в задачах анализа квазиоднородной Марковской модели сложных систем.

Свойство 1. Внешняя частная монотонность (ВЧМ) модели по k -экспериментальным наблюдениям. Модель M_1 обладает свойством 1, если $k_1^{M_1}, k_2^{M_1}$ — кратность повторов в плане экспериментов $\Pi(n, l, k, \theta)$ модели M_1 и с учетом неравенства $k_1^{M_1} > k_2^{M_1}$ выполняется $\Delta\gamma_{k_1}^{M_1} > \Delta\gamma_{k_2}^{M_1}$, следовательно, справедливо:

$$\Delta\gamma_{k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \quad (5)$$

и модель M_1 будем считать k — статистически устойчивой ($\Delta\gamma_k^{M_i}$) и монотонной по k — запуском модели (k — ВЧМ).

Свойство ВЧМ позволяет сравнивать модели в смысле эффективности оценок. Согласно определению плана экспериментов

свойство 1 может быть сформулировано для (n, l) параметров ММ. Модель M_1 обладает свойством 1, если с учетом неравенства $n_1^{M_1} > n_2^{M_1}$ или неравенства $l_1^{M_1} > l_2^{M_1}$ выполняется $\Delta\gamma_{n_1}^{M_1} > \Delta\gamma_{n_2}^{M_1}$ или $\Delta\gamma_{l_1}^{M_1} > \Delta\gamma_{l_2}^{M_1}$, следовательно, справедливо:

$$\Delta\gamma_{n_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{n_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \quad (6)$$

$$\Delta\gamma_{l_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{l_2}^{M_1}(\Pi_j, \theta_1^{M_1}). \quad (7)$$

Причем N^{M_1} и L^M определяются возможностями программных средств моделирования. Свойство ВЧМ (6) описывает (n — ВЧМ), для (7)-(l — ВЧМ) соответственно.

Для Марковской модели M_1 , реализованной по схеме С, предполагается исследование влияния длины цепи Маркова l^{M_1} , числа сгенерированных цепочек ММ n^{M_1} , а также числа запусков модели k^{M_1} на качество модели с точки зрения точности ε и оценки доверительного интервала p .

Таким образом, свойством 1 (выражения (5)-(7)) формулируют условия частной монотонности модели по n, l, k — экспериментальным данным. Сформулируем следствие 1 из свойства 1.

Следствие из свойства 1. Внешняя полная монотонность (ВПМ) модели. Пусть $k_1^{M_1}, k_2^{M_1}$ — кратность запусков в плане экспериментов $\Pi(n, l, k, \theta)$ модели M_1 , $n_i \geq n_0$, $l_i \geq l_0$. С учетом неравенства $k_1^{M_1} > k_2^{M_1}$ и с учетом неравенств $n_1^{M_1} > n_2^{M_1}$, $l_1^{M_1} > l_2^{M_1}$ выполняется $\Delta\gamma_{n_1}^{M_1} > \Delta\gamma_{n_2}^{M_1}$ и $\Delta\gamma_{l_1}^{M_1} > \Delta\gamma_{l_2}^{M_1}$ соответственно, а также выполняется $\Delta\gamma_{k_1}^{M_1} > \Delta\gamma_{k_2}^{M_1}$, следовательно справедливо:

$$\begin{aligned} \Delta\gamma_{n_1, l_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) &> \Delta\gamma_{n_1, l_1, k_2}^{M_1}(\Pi_j, \theta_1^{M_1}); \\ \Delta\gamma_{n_1, l_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) &> \Delta\gamma_{n_2, l_1, k_2}^{M_1}(\Pi_j, \theta_1^{M_1}); \\ \Delta\gamma_{n_1, l_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) &> \Delta\gamma_{n_1, l_2, k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \end{aligned} \quad (8)$$

и модель M_1 будем считать полностью монотонной по n, l, k – экспериментальным наблюдениям (обладающей ВПМ).

Аналогичным образом сформулируем свойство СУ по плану $\Pi(n, l, k, p)$ для доверительных вероятностей оценивания параметров модели.

Свойство 2. Монотонность модели по доверительной вероятности оценивания параметров. Модель M_1 обладает свойством 2, если сформулировано условие для определения доверительной вероятности p оценивания параметров модели, то для схемы сравнения мер СУ $\Delta\gamma_{p_1}^{M_1} > \Delta\gamma_{p_2}^{M_1}$ справедливо:

$$\Delta\gamma_{p_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{p_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } 0 < p^{M_1} < 1. \quad (9)$$

Свойство 2 может быть распространено на случай, когда сравниваются модели для одинаковых доверительных вероятностей $p_1^{M_1}$, но при $k_1^{M_1} > k_2^{M_1}$. Тогда (9) будет иметь вид:

$$\Delta\gamma_{p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{p_1, k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } 0 < p^{M_1} < 1. \quad (10)$$

В исследовании рассматривается в качестве примера $p_1^{M_1}$ для математического ожидания \tilde{m}_x .

Свойство 3. Монотонность модели по точности. Модель M_1 обладает свойством 3, если $\varepsilon_1^{M_1}, \varepsilon_2^{M_1}$ — точность результатов экспериментов в плане $\Pi(n, l, k, \theta)$ модели M_1 и с учетом неравенства $\varepsilon_1^{M_1} > \varepsilon_2^{M_1}$ выполняется $\Delta\gamma_{\varepsilon_1}^{M_1} > \Delta\gamma_{\varepsilon_2}^{M_1}$, следовательно, справедливо:

$$\Delta\gamma_{\varepsilon_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\varepsilon_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } 0 < \varepsilon_1^{M_1} < \varepsilon_0^{M_1}, \quad (11)$$

где $\varepsilon_0^{M_1}$ — критическое значение точности для модели M_1 .

Свойство 4 может быть распространено на два случая, когда сравниваются модели для одинаковых $p_1^{M_1}$, но при $k_1^{M_1} > k_2^{M_1}$; и на случай, когда для фиксированного k^{M_1} исследуются различные p^{M_1} . Тогда (11), может быть, будет распространено на два случая.

Случай 3.1. Монотонность модели по точности для фиксированного значения доверительной вероятности p^{M_1} :

$$\Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\varepsilon_1, p_1, k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } 0 < \varepsilon_1^{M_1} < \varepsilon_0^{M_1},$$

$$0 < p^{M_1} < 1. \quad (12)$$

Случай 3.2. Монотонность модели по точности для фиксированного значения числа запусков модели k^{M_1} :

$$\Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\varepsilon_1, p_2, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } 0 < \varepsilon_1^{M_1} < \varepsilon_0^{M_1},$$

$$0 < p^{M_1} < 1. \quad (13)$$

Вариативный характер погрешности ε формирует свойство модели при δ — отклонении числа репликаций модели оставаться в пределах заданного диапазона: $\varepsilon_{\min} < \varepsilon < \varepsilon_{\max}$. При этом варианты δ -отклонений числа запусков модели, например: $10 \pm \delta k$; $\delta = k$; $20 \pm \delta k$, могут быть определены как макроуровни анализа систем: $10 \pm \delta k$ — незначительное отклонение числа запусков модели; $10\delta k$ — десятикратное отклонение числа запусков модели; $\delta = k$ — фиксированное число запусков модели. Таким образом, сформулируем свойство, связанное с инвариантностью оценок качества модели к отклонению числа запусков (при больших объемах данных).

Свойство 4. Классификация оценок моделей по числу их запусков (расширение свойства 1). Свойство 1, отражающее сравнимость оценок СУ при $k_1^{M_1} > k_2^{M_1}$ для $\Delta\gamma_{k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{k_2}^{M_1}(\Pi_j, \theta_1^{M_1})$, на больших пространствах параметров МСИ и планах $\Pi(n, l, k, \theta)$ может быть сформулировано как ряд свойств моделей и планов:

Свойство 4.1. δ -мультипликативная ВЧМ модели по k -экспериментальным наблюдениям:

$$\Delta\gamma_{\delta k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\delta k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } k_1^{M_1}, k_2^{M_1} \in K^{M_1}, \delta \leq \tilde{\delta}_0. \quad (14)$$

Свойство 4.2. δ -аддитивная ВЧМ модели по k — экспериментальным наблюдениям:

$$\Delta\gamma_{\delta\pm k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\delta\pm k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } k_1^{M_1}, k_2^{M_1} \in K^{M_1}, \delta \leq \tilde{\delta}_0. \quad (15)$$

Свойство 4.3. δ — 10% -мультипликативная ВЧМ модели по k — экспериментальным наблюдениям:

$$\gamma_{10\delta k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \gamma_{10\delta k_2}^{M_1}(\Pi_j, \theta_1^{M_1}), \text{ при } k_1^{M_1}, k_2^{M_1} \in K^{M_1}, \delta \leq \tilde{\delta}_0. \quad (16)$$

Исходя из особенностей моделирования, ЛПР может предложить некоторые другие δ — свойства СУ моделей.

Определение 2. Модель M_x с учетом свойства 4, соответственно (14)-(16), считается (δk) -, $(\delta \pm k)$ или $(10\delta k)$ – статистически устойчивой и обозначается как $\Delta\gamma_{\delta k}^{M_x}, \Delta\gamma_{(\delta\pm k)}^{M_x}, \Delta\gamma_{10\delta k}^{M_x}$.

Например, при 1000 реализациях отклонение объема выборки на $\delta = 5$ по мнению ЛПР не вносит значительную погрешность в статистику и модель M_1 может быть $(\delta \pm k = 5 \pm 1000)$ — устойчивой $\Delta\gamma_{(\delta\pm k)}^{M_x}$.

Особенностью представленного исследования, в том числе является и то, что СТО рассматривается в рамках квазистационарного режима, при котором стационарность может наблюдаться не на всем исследуемом временном отрезке, а на ограниченных его участках, что позволит учесть особенности функционирования СТО.

Актуальность постановки задачи связана с вопросами практического характера, когда результаты анализа функционирования СТО оказываются нерепрезентативными и недостоверными, либо для достижения положительных исходов моделирования требуется избыток ресурсов. Четкая методологическая оценка макропоказателей и рычаги управления ресурсами на основе снижения размерности данных, варьирования точности, снижения статистической неопределенности, безусловно, важны как для развития методик МСИ, так и принятия решений в разных областях системного анализа.

4. Постановка задачи и описание модели. Пусть R — вектор рабочих режимов СТК; R_N — вектор номинального режима СТО, N — конечное множество номинальных режимов. Примем $|R_N - R_t| < \varepsilon$ как меру отклонения вектора текущего режима от номинального, причем это расстояние не превышает некоторую заданную величину $\varepsilon > 0$. Таким образом, R_t характеризует мгновенный «снимок» текущего режима СТО.

При статистических испытаниях для достижения наилучшего результата предполагается целесообразным проведение достаточного количества запусков модели [23-24]. Возникает необходимость определения достаточности выборки. Таким образом, рассматривается множество векторов $R_N^i, i = \overline{1, I}$, где I — количество реализаций исследуемого процесса. ММ строится по *schema A*: $\|P_{ij}\| \rightarrow \|P_{ij}^{Mm}\|$ в следующем фазовом пространстве состояний системы: $S = \{S_0, S_1, S_2, S_3, S_4\}$, где S_0 — рабочее состояние, S_1 — состояние отказа, S_2 — состояние восстановления системы после отказа, S_3 — состояние профилактики (в модели не фиксируется, работает или простаивает система в момент профилактики), S_4 — состояние сбоя. Это состояние, из которого может быть осуществлен переход в состояние отказа, после чего требуется восстановление, либо в начальное состояние без восстановления. Таким образом, S_4 — это состояние, при котором восстановление системы не требуется.

В модели присутствуют виртуальные переходы вида S_{ii} , исходя из того, что модель строилась в предположении, что $P_{ij} \approx 1/l$ где $l = \overline{1, L}$, L — полное множество переходов ММ [25]. На рисунке 1 приведен граф переходов сложного технического объекта, в таблице 1 — исходные вероятности переходов для имитационного моделирования.

Предложенный пример носит универсальный характер и показателен для анализа различных аспектов функционирования СТО, например принятия решений в области учащения профилактических работ или увеличения ресурса на восстановление СТО после отказа.

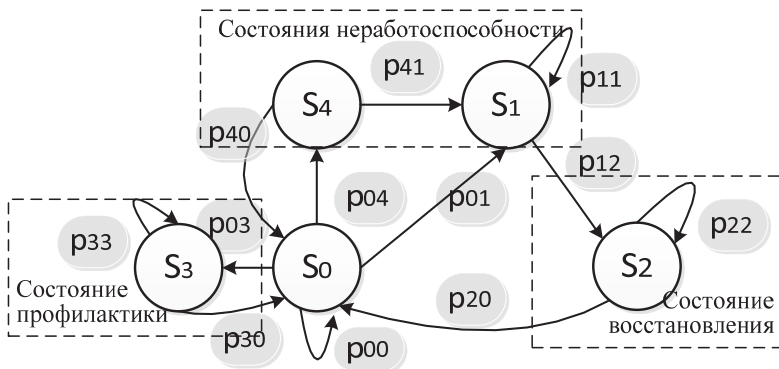


Рис. 1. Граф переходов в состояния системы со сбоями, восстановлениями и профилактикой

Зададим для примера значения исходной матрицы вероятностей переходов $\|P_{ij}\|$ с учетом $P_{ij} \approx 1/l$:

$$P = \begin{pmatrix} 0.35 & 0.25 & 0 & 0.20 & 0.20 \\ 0 & 0.50 & 0.50 & 0 & 0 \\ 0.55 & 0 & 0.45 & 0 & 0 \\ 0.55 & 0 & 0 & 0.45 & 0 \\ 0.55 & 0.45 & 0 & 0 & 0 \end{pmatrix}. \quad (17)$$

Для анализа вида переходов в матрице $\|P_{ij}\|$ введена мера дивергенции данных как соотношение:

$$\text{div}\|P_{ij}\| = \frac{\min P_{ij}}{\max P_{ij}}. \quad (18)$$

Выражение (18) отражает дивергенцию значений исходной матрицы переходов $\|P_{ij}\|$. В некоторых случаях целесообразно применять центрированный показатель $\text{div}\|P_{ij}\|$:

$$\text{div}\|P_{ij}^{M_i}\|_c = \frac{\left(\frac{P_{ij}}{L}\right) - \min(P_{ij}^{M_i})}{\max(P_{ij}^{M_i}) - \min(P_{ij}^{M_i})}. \quad (19)$$

Эта мера не обладает полнотой описания и используется в качестве иллюстративной характеристики моделей определенного свойства с малым разбросом значений переходных вероятностей: $P_{ij} \approx 1/l$, где $l = \overline{1, L}$, L — полное множество переходов ММ. Мера $\text{div}\|P_{ij}^{M_x}\|_c$ может быть полезна при исследовании моделей с высокой чувствительностью обнаружения исследуемых свойств объектов.

На рисунке 2 приведены зависимости сравнения дивергенции $\|P_{ij}^{M_i}\|$ для планов $\Pi_1(1..110, 1..100, \theta)$ и $\Pi_2(1..110, 1..200, \theta)$ с учетом свойства $1, l$ — ВЧМ.

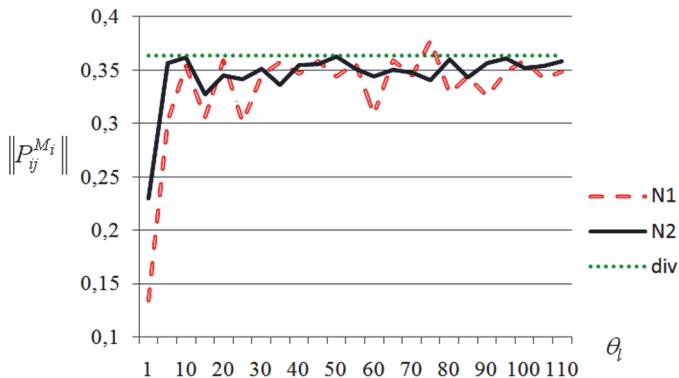


Рис. 2. Дивергенция значений матрицы $\|P_{ij}^{M_i}\|$ в зависимости от кратности сгенерированных цепочек ММ: $N1=100$, $N2=200$, div – дивергенция исходной матрицы

На рисунке 3 приведены зависимости сравнения дивергенции $div\|P_{ij}^{M_x}\|_c$ для планов $\Pi_1(1..110,1..100,\theta)$ и $\Pi_2(1..110,1..200,\theta)$ с учетом свойства 1, l — ВЧМ. Визуализация значений $div\|P_{ij}^{M_x}\|$ на рисунке 2, как и центрированная величина, рисунок 3 отражает процесс, сходящийся к исходной $div\|P_{ij}\|$ (верхняя точечная линия), начиная с $n=15$, причем для кратности реплицированных переходов $n=200$ сходимость более заметна, чем для $n=100$ (пунктирная линия).

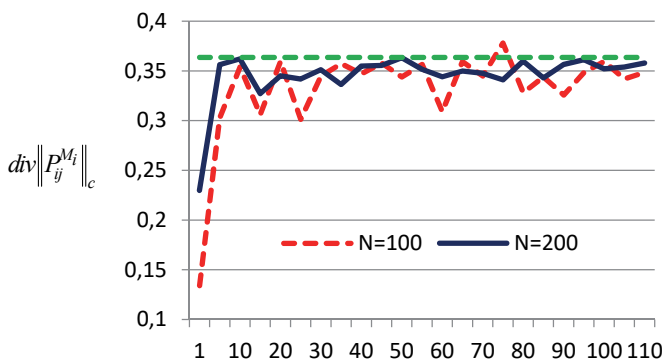


Рис. 3. Дивергенция значений матрицы $div\|P_{ij}^{M_i}\|_c$ в зависимости от кратности сгенерированных цепочек ММ

Для исследования СУ M_1 , представленной на рисунке 1, будем использовать следующие некоторые планы экспериментов, удовлетворяющих свойствам 1-4, по выражениям (5) – (16).

1. План $\Pi(1..100, 1..110, \theta)$, $\theta = \varepsilon$,
 $\Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{\varepsilon_1, p_2, k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$, свойство 3, определения (11), (13)

для схемы модели А.

2. Планы $\Pi_1(1..100, 1..110, \theta)$, $\Pi_2(1..200, 1..110, \theta)$ для схемы модели А, $\Delta\gamma_{n_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{n_2}^{M_1}(\Pi_j, \theta_1^{M_1})$, свойство 1, ВЧМ, определение (6).

3. План имитационного эксперимента $\Pi_1(50, 100, \theta)$, где $\theta = k = 20$, $k_1^{M_1} > k_2^{M_1}$, $\Delta\gamma_{k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{k_2}^{M_1}(\Pi_j, \theta_1^{M_1})$, свойство 1, ВЧМ, определение (5) для схемы модели А.

4. Планы $\Pi_1(1..100, 1..110, \theta)$, $\Pi_1(1..100, 1..110, \theta)$, для схемы модели А, $\Delta\gamma_{p_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{p_2}^{M_1}(\Pi_j, \theta_1^{M_1})$, свойство 2, определение (9) и $\Delta\gamma_{p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) > \Delta\gamma_{p_1, k_2}^{M_1}(\Pi_j, \theta_1^{M_1})$, свойство 2, определение (10) с уточнением $p_1^{M_1}$ на основе показателя статистической волатильности $\theta = \vartheta$.

5. Реализация имитационной модели. Имитационное моделирование проводилось с помощью созданного веб-ориентированного программного комплекса ПКAM, состоящего из трех модулей: модуль 1 — реализация процесса случайного блуждания по заданной матрице переходов $\|P_{ij}\|$ (схема P_v), модуль 2 (схема P_m), модуль 3 (схема А). Для визуализации используется графическая библиотека graphviz. Динамическое взаимодействие с пользователем реализовано средствами фреймворка Vue.js. Генератор случайных чисел используется встроенный криптостойкий с равномерным распределением значений от 0 до 1. Ресурсоёмкие операции выполняются асинхронно, что позволяет минимально затрагивать взаимодействие с интерфейсом. ПКAM имеет функции получения данных для статистических оценок параметров функционирования СТО различных размерностей и видов. На рисунке 4 приведен элемент интерфейса ПКAM с результатами восстановления модели с кратностью реплицированных переходов $n = 200$.

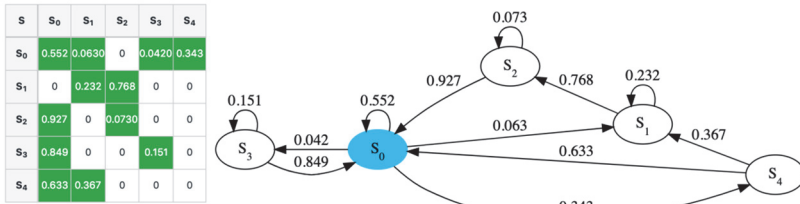


Рис. 4. Элемент интерфейса имитационного комплекса ПККАМ

Результаты имитации процесса функционирования СТО отражаются в виде одной или нескольких матриц $\|P_{ij}^{Mi}\|$, значения P_{ij} закрашиваются в зеленый или красный цвета при ϵ -различиях $\|P_{ij}^{Mi}\|$, параметры моделирования на рисунке 5: $k = 20, n = 200, l = 10$.

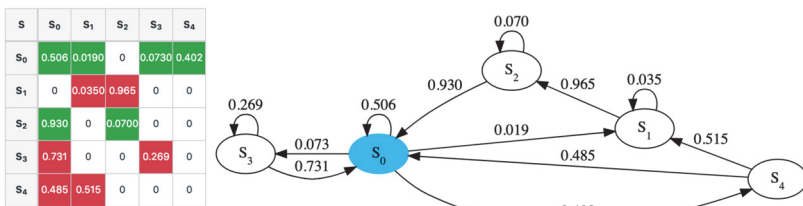


Рис. 5. Восстановление модели с кратностью сгенерированных цепочек ММ $n = 50$

ПККАМ имеет возможность регулировать точность вычислений, что позволяет оценить требуемые характеристики системы и условия ее стабильного функционирования. Таким образом, регулируя точность вычисления, можно изменить допустимые среднеквадратичные отклонения, принимать решения относительно полученных результатов в виде $\|\theta\|$ и дальнейшего их анализа в форме $\Delta\gamma_{\epsilon_1, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$.

6. Результаты анализа данных имитационного моделирования. Получение данных для анализа проходило по агрегированной схеме A . На основе свойства 2, рассмотрим диапазон: $\epsilon \in [0.01, 0.015, 0.02, 0.025, 0.03]$. На рисунке 6 приведены графики, отражающие результаты имитационного моделирования процесса случайного блуждания, состоящие в оценке количества ϵ -различных

элементов матрицы переходов $\|P_{ij}^{M_i}\|$ относительно $\|P_{ij}\|$, то есть отображены $\Delta\gamma_{\varepsilon, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$ для плана $\Pi(1..100, 1..110, \theta)$, где $\theta = \varepsilon$.

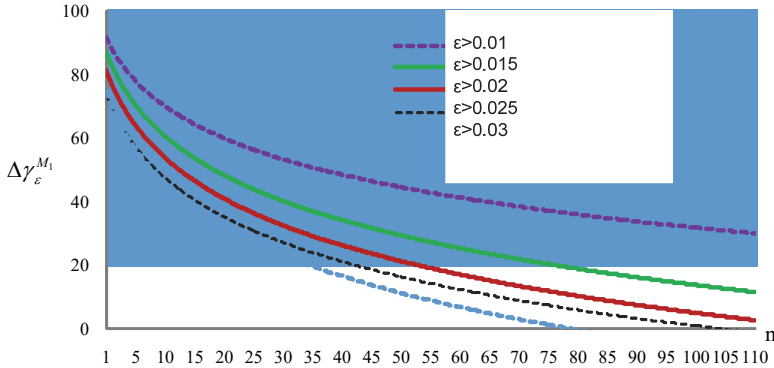


Рис. 6. Аппроксимирующие кривые для $\Delta\gamma_{\varepsilon}^{M_1}$ для план $\Pi(1..100, 1..110, \theta)$ для схемы A при различных пороговых значениях ε

На графиках показаны аппроксимирующие кривые на основе логарифмического тренда, построенного по $\Delta\gamma_{\varepsilon, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$.

Средняя величина достоверности построенных аппроксимирующих кривых для $\Delta\gamma_{\varepsilon, p_1, k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$ составляет $R^2 = 0.622$. Можно видеть, что при увеличении порога точности требование к числу переходов возрастает. Так для $\varepsilon = 0.01$ (верхняя кривая) значение $\Delta\gamma_{\varepsilon}^{M_1}$ для $n = 110$ находится на уровне 0.3, а для $\varepsilon = 0.03$ (нижняя кривая) значение $\Delta\gamma_{\varepsilon}^{M_1} = 0.3$ достигается при $n = 23$.

Для комплексной оценки однородности модели найдем среднее значение доли не совпавших значений в модели по схеме A . На рисунке 7 показаны вероятности ε -различных элементов матрицы $\|P_{ij}^{M_m}\|$ для $\Pi_1(1..100, 1..110, \theta)$ и $\Pi_2(1..200, 1..110, \theta)$ (св. 1).

Приведенные на рисунке 7 кривые отражают снижение меры $\Delta\gamma_{\varepsilon, p_1, n}^{M_1}$ неустойчивости модели при увеличении ε для $n = 200$

относительно $n=100$. Так, для $\varepsilon=0.03$ (верхняя кривая) $\Delta\gamma_{\varepsilon, P_1, 100}^{M_1} \approx 0.18$, (нижняя кривая) $\Delta\gamma_{\varepsilon, P_1, 200}^{M_1} \approx 0.1$.

Определение 3. Мера неустойчивости процедуры восстановления матрицы переходов как модельный параметр (ε -различие $\|P_{ij}^{M_i}\|$) по схеме A определяется как *статистическая волатильность* на множестве из n -реализаций с учетом стандартного отклонения s : $\mathcal{G} = s \cdot \sqrt{n}$.

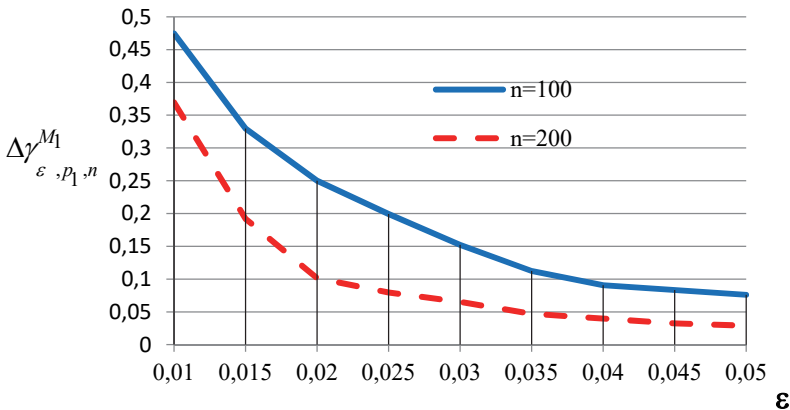


Рис.7. Эмпирические оценочные кривые $\Delta\gamma_{\varepsilon, P_1, n}^{M_1}$ по всем переходам

при ε -различии $\|P_{ij}\|$ и $\|P_{ij}^{M_m}\|$

Выражение для \mathcal{G} позволяет по аналогии с финансовой волатильностью [21] оценить степень неустойчивости статистик (нестационарности статистик) в зависимости от объема повторных запусков модели. Для моделей больших размерностей и переменным числом запусков этот показатель позволит получить числовую меру верификации модели с точки зрения устойчивости (неустойчивости) статистик.

На рисунке 8 представлены оценки статистической волатильности ММ, построенные по планам: $\Pi_1(1..100, 1..110, \theta)$ и $\Pi_2(1..200, 1..110, \theta)$, где $\theta = \mathcal{G}$. То есть модель рассматривается в диапазоне параметров $n_1 = 1..100, n_2 = 1..200, l = 1..110$.

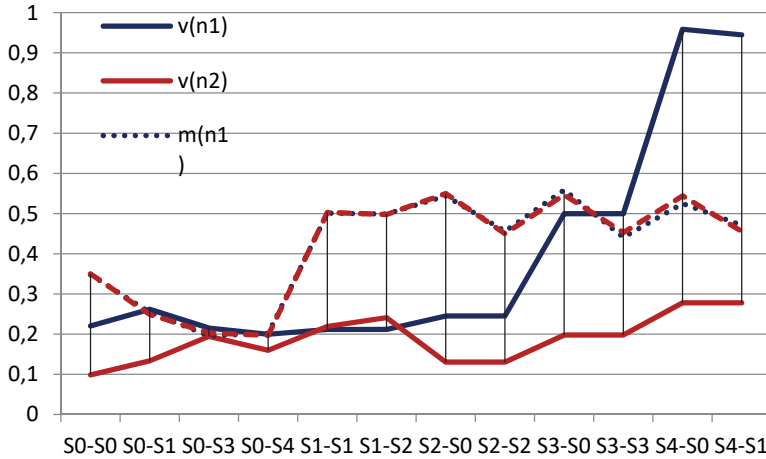


Рис. 8. Оценки величины статистической волатильности ММ, построенные для планов $\Pi_1(1..110, 1..100, \theta)$ и $\Pi_2(1..110, 1..200, \theta)$ при $n1=100, n2=200$

Можно отметить существенное повышение неустойчивости модели (волатильности) относительно математического ожидания при уменьшении кратности запусков модели в два раза.

Для оценки доверительного интервала в простейшем случае используем статистику Стьюдента

$$t = S_{n-1} = \frac{\Gamma(n/2)}{\sqrt{(n-1)\pi \cdot \Gamma((n-1)/2)}} \left(1 + \frac{t^2}{n-1}\right)^{n/2}. \text{ С учетом ввода переменных}$$

$$T = \frac{\tilde{m}_x - m_x}{\sqrt{\tilde{D}/n}} \in S_{n-1}, \text{ где } S_{n-1} \text{ — распределение по закону Стьюдента;}$$

\tilde{m}_x — оценка математического ожидания; \tilde{D} — оценка дисперсии распределения, получим:

$$\frac{\varepsilon}{\sigma_{\tilde{m}_x} \cdot \sqrt{2}} = S^{-1}(\beta), \quad \varepsilon = \sqrt{\frac{\tilde{D}}{n}} S^{-1}(\beta). \quad (20)$$

Таким образом, для планов $\Pi_1(1..100, 1..110, \theta)$ и $\Pi_2(1..200, 1..110, \theta)$ получим для вероятностей $\beta_1 = 0,9$ и $\beta_2 = 0,99$ соответственно: $S^{-1}(\beta_1) = 1,65$; $S^{-1}(\beta_2) = 2,61$. Тогда $\varepsilon_1 = 0,021$,

$\varepsilon_2 = 0,033$, отсюда получим интервалы оценки для вероятности перехода: $p_{00} (\tilde{m}_x = 0,35)$: $\gamma_1 = 0,33$; $\gamma_2 = 0,37$ для $\beta_1 = 0,9$ и $\gamma_1 = 0,31$; $\gamma_2 = 0,38$ для $\beta_2 = 0,99$ с учетом свойства 2. На рисунке 9 отражены оценки доверительных интервалов для \tilde{m}_x , построенные для модели M_x по плану: $\Pi_1(50,100,\theta)$, где $\theta_i : \langle k_i, \tilde{m}_x, p_i, v_i \rangle$; $k = 20$, v_i — статистическая волатильность модели.



Рис. 9. Вероятность и доверительные интервалы для оценки \tilde{m}_x вероятности по $\Pi_1(50,100,20)$ с учетом v_i

Вероятность p , с которой оцениваются доверительные интервалы СВ в приведенной схеме, зависят от размера выборки N . Очевидно, что чем больше N , тем доверительная вероятность p выше. Но не для всех инженерных задач возможно получить репрезентативную выборку, в связи с чем целесообразно применять уточненные доверительные интервалы. На рисунке 10 приведена гистограмма распределения оценки СВ $\|P_{ij}^{Mm}\|$ (переход $S_0 \rightarrow S_0$) и доверительные интервалы для оценки \tilde{m}_x для переходных вероятностей (вертикальные пунктирные линии) с учетом показателя волатильности модели (вертикальные отрезки). Очевидно, распределение оценки СВ несимметрично, имеет следующие статистические оценки: асимметрия: $-0,32$, эксцесс: $-0,39$, медиана: $0,355$, мода: $0,347$.

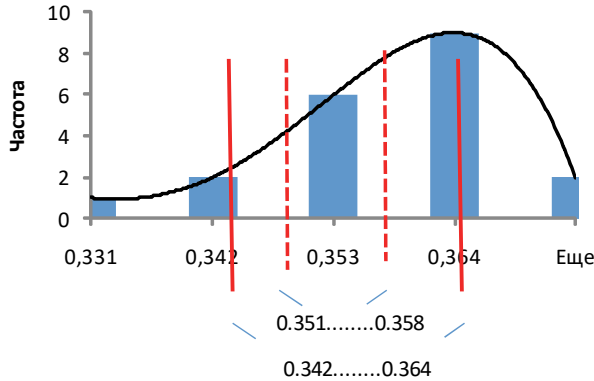


Рис. 10. Гистограмма распределения оценки СВ вероятности перехода $S_0 \rightarrow S_0$ и доверительные интервалы для $\Pi_1(50,100,20)$

На рисунке 11 приведена гистограмма распределения оценки случайной величины (СВ) $\|P_{ij}^{Mm}\|$ (переход $S_0 \rightarrow S_4$). В отличие от ранее рассмотренного виртуального перехода $S_0 \rightarrow S_0$ распределение оценки СВ более симметрично (асимметрия: -0,15, эксцесс: -0,67, медиана: 0,195, мода: 0,199).

С учетом статистической волатильности модели для перехода $S_0 \rightarrow S_0$ доверительный интервал изменился с 0,0035 до 0,0091 (рисунки 6 и 7); для перехода $S_0 \rightarrow S_4$ доверительный интервал изменился с 0,003 до 0,0064 (рисунок 8).

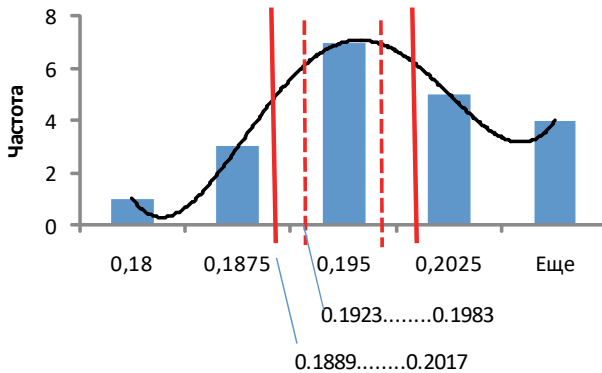


Рис. 11. Гистограмма распределения оценки СВ вероятности перехода $S_0 \rightarrow S_4$ и доверительный интервал для $\Pi_1(50,100,20)$

Таким образом, согласно случаю 3.2 (свойство 3) $\Delta\gamma_{\varepsilon_1 \cdot p_1 \cdot k_1}^{M_1}(\Pi_j, \theta_1^{M_1}) < \Delta\gamma_{\varepsilon_1 \cdot p_2 \cdot k_1}^{M_1}(\Pi_j, \theta_1^{M_1})$, где $p_1^{M_1} \pm \nu \rightarrow p_2^{M_1}$. Это показывает, что существующие статистические оценки в определенных случаях вносят существенную погрешность в анализе качества моделей функционирования СТО.

На рисунке 12 приведена гистограмма распределения оценки $C \|P_{ij}^{M_m}\|$ (переход $S_0 \rightarrow S_4$) для плана $\Pi_1(60, 100, 20)$ с учетом свойства 4.

Статистическая характеристика распределения СВ симметрична (асимметрия: 0,59, эксцесс: 1,36, медиана: 0,2215, мода: 0,223). В планах $\Pi_1(60, 100, 20)$ относительно $\Pi_2(50, 100, 20)$, согласно свойству 4.2, имеет место ситуация δ -аддитивности ВЧМ модели с параметрами $\delta = 10$, $k = 60$ и показателями волатильности и СКО для этого случая: $\nu_1 = 0.064, s = 0.0014$ и $\nu_2 = 0.023, s = 0.0005$ соответственно.

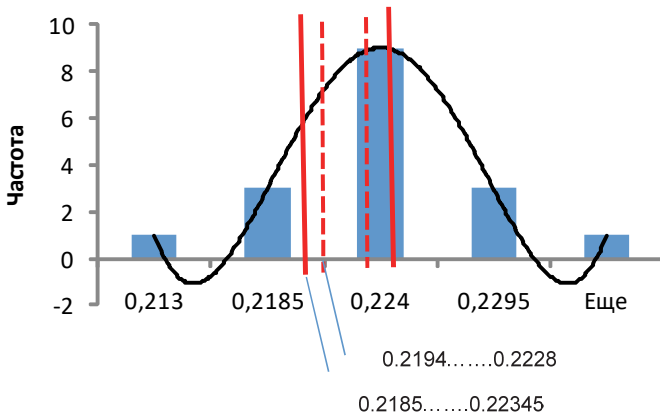


Рис. 12. Гистограмма распределения оценки СВ вероятности перехода $S_0 \rightarrow S_4$ и доверительный интервал для $\Pi_1(60, 100, 20)$

Анализ модели, приведенной на рисунке 1, показал, что для принятия решений по процессу функционирования СТО целесообразно иметь рычаги управления режимами в рамках заданных условий $|R_N - R_r| < \varepsilon$. На рисунке 13 изображены результаты моделирования по плану $\Pi(100, 15..70, 9)$ для перехода $S_0 \rightarrow S_1$ (переход в состояние отказа) с учетом рассчитанных доверительных интервалов.

Отметим, что для промежутка $l = 35..55$ фактические значения вероятности перехода выходят за пределы рассчитанных стандартным способом интервалов доверия, тогда как учет статистической волатильности модели (красная кривая) уточнил ожидаемый показатель.

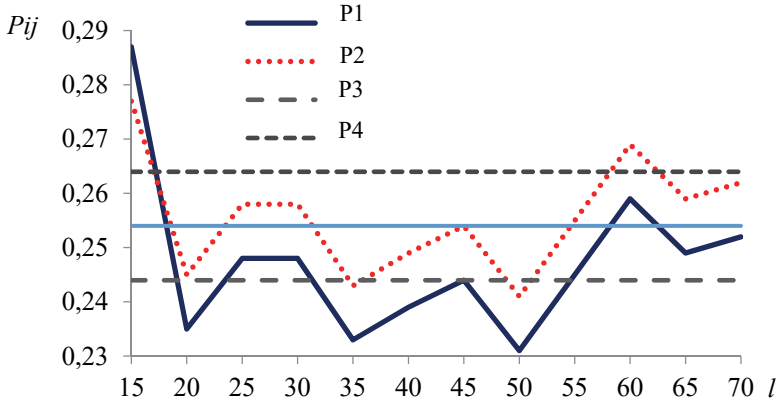


Рис. 13. Результаты моделирования $\|P_{ij}^{M,m}\|$ по плану П(100,15..70,9) для перехода $S_0 \rightarrow S_1$ с учетом доверительных интервалов и оценки статистической волатильности: p1 – при $n=100$, p2 – с учетом волатильности модели, p3 – доверительная граница (max), p4 – доверительная граница (min)

При наличии у ЛПР требований к точности оценке показателей режимов СТО $|R_N - R_t| < \varepsilon$ в диапазоне $\varepsilon = 0.01..0.025$ и ресурса, позволяющего проводить не более $k = 20$ запусков модели при изменении числа сгенерированных цепочек ММ от 100 до 200, на рисунке 14 схематически показан принцип, позволяющий принимать решения по дополнительному проведению профилактических работ.

Для $\varepsilon = 0.015$ оценка СУ по мере $\Delta\gamma_{\varepsilon_1, P_1, k_1}^{M_1}$ колеблется от 0,67 до 0,81, что весьма существенно при принятии решений о наличии проблемной ситуации при исследовании перехода в состояние отказа. Если требования к ресурсам не позволяют проводить дополнительные запуски модели в увеличить число переходов, то модель может считаться $\Delta\gamma_{\varepsilon_1, P_1, k_1}^{M_1}$ устойчивой при $\varepsilon = 0.015, n = 100, k = 20$.

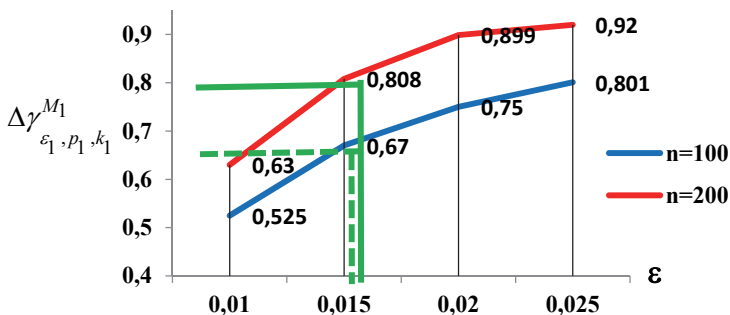


Рис. 14. Оценки СУ $\Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1}$ при изменении числа реплицированных переходов

Ввиду того, что для проверки гипотез о статистической устойчивости и статистической волатильности моделей требуется довольно большой объем расчетных процедур, целесообразна автоматизация этого процесса в форме системы поддержки принятия решений, что является темой для дальнейших исследований авторов.

7. Заключение. Исследование статистической устойчивости в задачах анализа квазиоднородной модели сложных систем в частности и процессов функционирования СТО в целом, особенно на границе изменения этапов их жизненного цикла, например при переходе к режиму деградиционных изменений, режиму устаревания или, напротив, совершенствования, сопряжено с рядом принципиальных особенностей, которые могут ограничить использование принятых статистик. Существующие исследования направлены в основном на изучение обширной области оценки качества моделей в рамках существующих мер.

В связи с многомерностью планов экспериментов, сложностью задач квалиметрического анализа моделей, рассмотрение статистической устойчивости в общем виде не представляется возможным. Перспективным является конструктивный подход к определению мер СУ. Такой подход был реализован для СТО. Приведенные в исследовании статистическая устойчивость и статистическая волатильность позволили получить информацию о качестве модели с точки зрения понимания устойчивости модели как фактора принятия решений в области анализа и синтеза систем. Предложенные формальные описания и свойства, характеризующие СУ моделей, могут быть полезны для специалистов в области анализа сложных технических объектов, в том числе в рамках Марковских моделей.

Вычисленные оценки для приведенного примера показали, что, например, для $\varepsilon = 0.015$ оценка СУ $\Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1}$ колеблется от 0,67 до 0,81 $\left(0 < \Delta\gamma_{\varepsilon_1, p_1, k_1}^{M_1} < 1\right)$ при различных размерностях выборок. Это существенно при принятии решений о наличии проблемной ситуации при исследовании перехода в состояние отказа.

С учетом введенных определений и свойств СУ в примере модель может считаться $\Delta\gamma_{\varepsilon_1, k_1, m_1}^{M_1}$ -устойчивой при $\varepsilon = 0.015, n = 100, k = 20$.

Дальнейшие исследования предполагается вести в направлении расширения группы мер валидации моделей, а также создания системы поддержки и принятия решений по рабочим режимам в области практического применения полученных результатов для исследования функционирования СТО, в том числе в критических условиях и на этапе деградации их свойств, а также в направлении учета временных параметров имитации процесса функционирования СТО на основе полумарковской модели.

Литература

1. *Herniter M.E.* Introduction to Model-Based System Design // Rose Hulman Institute of Technology. 2010. URL: http://wiki.ece.rosehulman.edu/herniter/images/8/8d/MBSD1_Lecture_Notes_Complete_Winter09-10.pdf (дата обращения: 20. 06. 2018).
2. *Panella I., Hardwick G.* Model Oriented System Design Applied to Commercial Aircraft Secondary Flight Control Systems // International Conference Simulation and Modeling Methodologies, Technologies and Applications. 2017. pp. 55–76.
3. *Боев В. Д.* Исследование адекватности GPSS World и AnyLogic при моделировании дискретно-событийных процессов: Монография // СПб: ВАС. 2011. 404 с.
4. *Скатков А.В., Балакирева И.А.* Обеспечение функциональной устойчивости операционных характеристик систем мониторинга окружающей среды при произвольном входном потоке данных // Системы контроля окружающей среды. 2017. № 8(28). С. 47–54.
5. *Laaksonen O., Peltoniemi M.* The essence of dynamic capabilities and their measurement // International Journal of Management Reviews. 2018. vol. 20(2). pp. 184–205.
6. *Langville A.N., Meyer C.D.* Updating Markov chains with an eye on Google's PaGeRank // SIAM Journal on Matrix Analysis and Applications. 2006. vol. 27(4). pp. 968–987.
7. *Мусаев А.А., Скворцов М.С.* Методы параметрической оптимизации надежности структур но-сложных технических систем // Труды СПИИРАН. 2008. Вып. 6. С. 44–50.
8. *Li Y.F., Zio E.* A multi-state model for the reliability assessment of a distributed generation system via universal generating function // Reliability Engineering & System Safety. 2012. vol. 106. pp. 28–36.

9. *Chen N., Majda A.* Conditional Gaussian systems for multiscale nonlinear stochastic systems: prediction, state estimation and uncertainty quantification // *Entropy*. 2018. vol. 20. no. 7. pp. 1–80.
10. *Кондрашков А.В., Пичугин Ю.А.* Идентификация и статистическая проверка устойчивости модели Вольтерры // *Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Физико-математические науки*. 2014. № 1(189). С. 124–135.
11. *Garza-Reyes J.A. et al.* A PDCA-based approach to Environmental Value Stream Mapping (E-VSM) // *Journal of Cleaner Production*. 2018. vol. 180. pp. 335–348.
12. *Gorban I.I.* The Statistical Stability Phenomenon // *Springer*. 2017. 361 p.
13. *Mishura I.S.* Stochastic Calculus for Fractal Brownian Motion and Related Processes // *Springer*. 2008. 393 p.
14. *Микони С.В., Соколов Б.В. Юсупов Р.М.* Квалиметрия моделей и полимодельных комплексов: монография // М.: РАН. 2018. 314 с.
15. *Sokolov B.V., Yusupov R.M.* Information Fusion Models' Quality Estimation And Models' Quality Control Theory // VI ISTC Scientific Advisory Committee Seminar Science and Computing. 2003. pp. 102–104.
16. *Ростовцев Ю.Г., Юсупов Р.М.* Проблема обеспечения адекватности субъектно-объектного моделирования // *Известия вузов. Приборостроение*. 1991. Т. 24. № 7. С. 7–14.
17. *Советов Б.Я. и др.* Имитационное моделирование систем // Петродворец: ВУНЦ ВМФ. 2010. 343 с.
18. *Dolgui A., Ivanov D., Sokolov B.* Scheduling of recovery action in the supply chain with resilience analysis consideration // *International Journal of Production Research*. 2018. vol. 56. no. 19. pp. 6473–6490.
19. *Zhao Y. et al.* Fast noisy image quality assessment based on free-energy principle // *Communications in Computer and Information Science*. 2018. vol. 815. pp. 290–299.
20. *Okhtilev M.Yu. et al.* Methods and Algorithms of Integrated Modeling of Complex Technical Objects in Dynamically Changing Conditions // *Proceedings of the International Scientific Conference MMET NW*. 2018. pp. 282–284.
21. *Sokolov B. et al.* Logic Dynamic Model And Algorithms Of Operation Complex // *European Modeling & Simulation Symposium (EMSS-2018)*. 2018. pp. 59–67.
22. *Degiannakis S., Floros C.* Methods of Volatility Estimation and Forecasting. In: *Modelling and Forecasting High Frequency Financial Data* // *Palgrave Macmillan*. 2015. pp. 58–109.
23. *Gerasimova D.S., Sayapin A.V., Palukhin A.A., Katsura A.V.* Application of the bootstrap method for statistical characteristics assessment of aircraft components' small samples // *Сибирский журнал науки и технологий*. 2018. № 3. С. 482–488. URL: <https://cyberleninka.ru/article/n/application-of-the-bootstrap-method-for-statistical-characteristics-assessment-of-aircraft-components-small-samples> (дата обращения: 24.03.2019).
24. *Park K.I.* Fundamentals of Probability and Stochastic Processes with Applications to Communications // *Springer International Publishing*. 2018. 273 p.
25. *Tien I., Der Kiureghian A.* Algorithms for bayesian network modeling and reliability assessment of infrastructure systems // *Reliability Engineering & System Safety*. 2016. vol. 156. pp. 134–147.

Доронина Юлия Валентиновна — д-р техн. наук, доцент, профессор, кафедра информационных систем института информационных технологий и управления в технических системах, Севастопольский государственный университет (СевГУ). Область научных интересов: системный анализ и синтез, реинжиниринг информационных систем, моделирование организационных процессов. Число научных публикаций — 100.

juvado@yandex.ru; ул. Университетская, 33, 299053, Севастополь, Российская Федерация;
р.т.: +7 (8692) 435-138.

Скатков Александр Владимирович — д-р техн. наук, профессор, профессор, кафедра информационных технологий и компьютерных систем института информационных технологий и управления в технических системах, Севастопольский государственный университет (СевГУ). Область научных интересов: системный анализ, критические инфраструктуры, гарантоспособные сервисы, распределенные системы. Число научных публикаций — 300. vm1945@mail.ru; ул. Университетская, 33, 299053, Севастополь, Российская Федерация; р.т.: +7 (8692) 435-173.

YU.V. DORONINA, A.V. SKATKOV
**STATISTICAL STABILITY ANALYSIS OF STATIONARY
MARKOV MODELS**

Doronina YuV., Skatkov A.V. Statistical Stability Analysis of Stationary Markov Models.

Abstract. An approach is proposed to assess the quality of stationary Markov models without absorbing states on the basis of a measure of statistical stability: the description is formulated and its properties are determined. It is shown that the estimates of statistical stability of models were raised by different authors, either as a methodological aspect of the model quality, or within the framework of other model properties. When solving practical problems of simulation, for example, based on Markov models, there is a pronounced problem of ensuring the dimension of the required samples. On the basis of the introduced formulations, a constructive approach to solving the problems of sample size optimization and statistical volatility analysis of the Markov model to the emerging anomalies with restrictions on the accuracy of the results is proposed, which ensures the required reliability and the exclusion of non-functional redundancy.

To analyze the type of transitions in the transition matrix, a measure of its divergence (normalized and centered) is introduced. This measure does not have the completeness of the description and is used as an illustrative characteristic of the models of a certain property. The estimation of the divergence of transition matrices can be useful in the study of models with high sensitivity of detection of the studied properties of objects. The key stages of the approach associated with the study of quasi-homogeneous models are formulated.

Quantitative estimates of statistical stability and statistical volatility of the model are proposed on the example of modeling a real technical object with failures, recovery and prevention. The effectiveness of the proposed approaches in solving the problem of statistical stability analysis in the problems of qualimetric analysis of quasi-homogeneous models of complex systems is shown. On the basis of the offered constructive approach the operational tool of decision-making on parametric and functional adjustment of difficult technical objects on long-term and short-term prospects is received.

Keywords: Statistical Stability, Quasi-Homogeneous Model, Statistical Volatility, Random Walk, Markov Chain, Complex Technical System, Accuracy, Transition Probability, Number of Process Implementations, Model Divergence.

Doronina Julia Valentinovna — Ph.D., Dr.Sci., Associate Professor, Professor, Department Information Systems of Institute of Information Technologies and Management in the Technical Systems, Sevastopol State University (SEVGU). Research interests: system analysis and synthesis, reengineering of information systems, modeling of organizational processes. The number of publications — 100. juvado@yandex.ru; 33, Universitetskaya str., 299053, Sevastopol, Russian Federation; office phone: +7 (8692) 435–138.

Skatkov Aleksander Vladimirovich — Ph.D., Dr.Sci., Professor, Professor, Department of information technologies and computer systems of Information Technologies and Management in the Technical Systems, Sevastopol state University (SevSU). Research interests: system analysis, critical infrastructure, dependable services, distributed systems. The number of publications — 300. vm1945@mail.ru; 33, Universitetskaya str., 299053, Sevastopol, Russian Federation; office phone: +7 (8692) 435-173.

References

1. Hermiter M.E. Introduction to Model-Based System Design // Rose Hulman Institute of Technology. 2010. Available at: <http://wiki.ece.rosehulman.edu/>

- hermiter/images/8/8d/MBSD1_ Lecture_ Notes_ Complete_ Winter09-10.pdf (accessed: 20.06.2018).
2. Panella L., Hardwick G. Model Oriented System Design Applied to Commercial Aircraft Secondary Flight Control Systems. International Conference Simulation and Modeling Methodologies, Technologies and Applications. 2017. pp. 55–76.
 3. Boev V.D. *Issledovanie adekvatnosti GPSS World i AnyLogic pri modelirovanii diskretno-sobytnykh processov: Monografiya* [Study of the adequacy of GPSS World and AnyLogic in the modeling of discrete-event processes]. SPb: VAS. 2011. 404 p. (In Russ.).
 4. Skatkov A.V., Balakireva I.A. [Ensuring functional stability of operating characteristics of environmental monitoring systems at arbitrary input data flow]. *Sistemy kontrolya okruzhayushhej sredy – Monitoring systems of environment*. 2017. vol. 8(28) pp. 47–54. (In Russ.).
 5. Laaksonen O., Peltoniemi M. The essence of dynamic capabilities and their measurement. *International Journal of Management Reviews*. 2018. vol. 20(2). pp. 184–205.
 6. Langville A.N., Meyer C.D. Updating Markov chains with an eye on Google's PaGeRank. *SIAM Journal on Matrix Analysis and Applications*. 2006. vol. 27(4). pp. 968–987.
 7. Musaev A.A., Skvorcov M.S. [Methods of parametric optimization of reliability of structurally complex technical systems]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2008. vol. 6. pp. 44–50. (In Russ.).
 8. Li Y.F., Zio E. A multi-state model for the reliability assessment of a distributed generation system via universal generating function. *Reliability Engineering & System Safety*. 2012. vol. 106. pp. 28–36.
 9. Chen N., Majda A. Conditional Gaussian systems for multiscale nonlinear stochastic systems: prediction, state estimation and uncertainty quantification. *Entropy*. 2018. vol. 20. no. 7. pp. 1–80.
 10. Kondrashkov A.V., Pichugin Ju.A. [Identification and statistical verification of Volterra model stability]. *Nauchno-tehnicheskie ведомosti SpbGPU. Fiziko-matematicheskie nauki – St. Petersburg Polytechnic University Journal. Physics and Mathematics*. 2014. vol. 1(189). pp. 124–135. (In Russ.).
 11. Garza-Reyes J.A. et al. A PDCA-based approach to Environmental Value Stream Mapping (E-VSM). *Journal of Cleaner Production*. 2018. vol. 180. pp. 335–348.
 12. Gorban I.I. *The Statistical Stability Phenomenon*. Springer. 2017. 361 p.
 13. Mishura I.S. *Stochastic Calculus for Fractional Brownian Motion and Related Processes*. Springer. 2008. 393 p.
 14. Mikoni S.V., Sokolov B.V., Jusupov R.M. *Kvalimetrija modelej i polimodel'nykh kompleksov: monografiya* [Qualimetry of models and polydivide complexes: monograph]. M.: RAN. 2018. 314 p. (In Russ.).
 15. Sokolov B.V., Yusupov R.M. Information Fusion Models' Quality Estimation And Models' Quality Control Theory. VI ISTC Scientific Advisory Committee Seminar Science and Computing. 2003. pp. 102–104.
 16. Rostovcev Ju.G., Jusupov R.M. [The problem of ensuring the adequacy of subject-object modeling]. *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie – Journal of Instrument Engineering*. 1991. Issue 24. vol. 7. pp. 7–14. (In Russ.).
 17. Sovetov B.Ja. et al. *Imitacionnoe modelirovanie sistem* [Simulation modeling of systems]. Petrodvorec: VUNC VMF. 2010. 343 p. (In Russ.).
 18. Dolgui A., Ivanov D., Sokolov B. Scheduling of recovery action in the supply chain with resilience analysis consideration. *International Journal of Production Research*. 2018. vol. 56. no. 19. pp. 6473–6490.
 19. Zhao Y. et al. Fast noisy image quality assessment based on free-energy principle. *Communications in Computer and Information Science*. 2018. vol. 815. pp. 290–299.
 20. Okhtilev M.Yu. et al. Methods and Algorithms of Integrated Modeling of Complex Technical Objects in Dynamically Changing Conditions. Proceedings of the International Scientific Conference MMET NW. 2018. pp. 282–284.

21. Sokolov B. et al. Logic Dynamic Model And Algorithms Of Operation Complex. European Modeling & Simulation Symposium (EMSS-2018). 2018. pp. 59–67.
22. Degiannakis S., Floros C. Methods of Volatility Estimation and Forecasting. In: Modelling and Forecasting High Frequency Financial Data. Palgrave Macmillan. 2015. pp. 58–109.
23. Gerasimova D.S., Sayapin A.V., Palukhin A.A., Katsura A.V. [Application of the bootstrap method for statistical characteristics assessment of aircraft components' small samples]. *Sibirskij zhurnal nauki i tekhnologii – Siberian Journal of Science and Technology*. 2018. vol. 3. pp. 482–488. Available at: <https://cyberleninka.ru/article/n/application-of-the-bootstrap-method-for-statistical-characteristics-assessment-of-aircraft-components-small-samples> (accessed: 24.03.2019).
24. Park K.I. Fundamentals of Probability and Stochastic Processes with Applications to Communications. Springer International Publishing. 2018. 273 p.
25. Tien I., Der Kiureghian A. Algorithms for bayesian network modeling and reliability assessment of infrastructure systems. *Reliability Engineering & System Safety*. 2016. vol. 156. pp. 134–147.

Ф.А. ТАУБИН, А.Н. ТРОФИМОВ
**КАСКАДНОЕ КОДИРОВАНИЕ ДЛЯ МНОГОУРОВНЕВОЙ
ФЛЭШ-ПАМЯТИ С ИСПРАВЛЕНИЕМ ОШИБОК МАЛОЙ
КРАТНОСТИ ВО ВНЕШНЕЙ СТУПЕНИ**

Таубин Ф.А., Трофимов А.Н. Каскадное кодирование для многоуровневой флэш-памяти с исправлением ошибок малой кратности во внешней ступени.

Аннотация. Один из эффективных подходов к организации помехоустойчивого кодирования в многоуровневой флэш-памяти связан с использованием каскадных конструкций на основе многомерных целочисленных решеток, используемых для построения внутреннего кода. Характерной особенностью таких каскадных конструкций является доминирование доли сложности внешнего декодера в общей сложности каскадного декодера. Учитывая, что в практических приложениях сложность декодирования, как правило, ключевое ограничение, определяющее возможность использования помехоустойчивого кодирования для многоуровневой флэш-памяти, каскадные конструкции со сравнительно малой сложностью внешнего декодера могут оказаться привлекательным решением в рамках обменного соотношения «плотность записи — сложность декодирования». Рассмотрена каскадная схема кодирования для многоуровневой флэш-памяти, в которой в качестве внутренней ступени используются коды на основе решеток Барнса — Уолла, а в качестве внешней ступени используется код Рида — Соломона с исправлением малого числа ошибок — не более 4...5.

Анализ помехоустойчивости предложенной каскадной схемы выполнен применительно к модели, отражающей основные физические особенности ячейки флэш-памяти с неравномерно расположенными целевыми уровнями напряжения в ячейке и дисперсией шума, зависящей от записанного значения (input-dependent additive Gaussian noise, ID-AGN). Для этой модели в работе развита модификация ранее предложенного авторами подхода к оценке вероятности ошибки декодирования внутреннего кода, основанная на использовании параллельной структуры кодовой решетки внутреннего кода, что позволяет существенно понизить сложность вычислений и ускорить получение окончательного результата. Приведены численные результаты, иллюстрирующие степень снижения достижимой плотности записи при введении ограничения на число исправляемых кодом Рида — Соломона ошибок — не более 4 — для широкого диапазона значений времени хранения данных и числа циклов перезаписи.

Ключевые слова: каскадный код, решетки Барнса — Уолла, код Рида — Соломона, многоуровневая флэш-память, вероятность ошибки декодирования, сложность декодирования.

1. Введение. Помехоустойчивое кодирование является составной частью технологии производства флэш-памяти. Простейшие варианты коррекции ошибок — коды Хэмминга — использовались исключительно для флэш-памяти с одноуровневыми ячейками (single-level cell, SLC). Увеличение плотности записи, достигаемое при переходе от флэш-памяти с одноуровневыми ячейками к многоуровневой (multi-level cell, MLC), потребовало использования более эффективных методов кодирования. В индустрии производства флэш-памяти стандартом выбора де-факто являются коды Боуза—Чоудхури — Хоквингема (БЧХ) [1-5]. Однако эффективность использования БЧХ-кодов ока-

зывается весьма чувствительной к исходной вероятности ошибки (raw bit error rate) в многоуровневой флэш-памяти. В частности, необходимые значения избыточности и сложности декодирования БЧХ-кодов, при которых может быть обеспечена требуемая вероятность ошибки порядка 10^{-12} ... 10^{-16} , быстро возрастают (часто до неприемлемых величин) при достижении исходной вероятности ошибки (raw bit error rate) уровня 10^{-4} ... 10^{-3} [6].

Это обстоятельство диктует необходимость поиска более эффективных подходов к организации введения помехоустойчивого кодирования в многоуровневой флэш-памяти. Один из таких подходов, часто рассматриваемый по множеству публикаций, состоит в использовании кодов с малой плотностью проверок на четность (LDPC, Low-Density Parity-Check), имеющих длину порядка несколько тысяч символов и декодируемых с использованием мягких решений (см. к примеру, работы [7-9]). Основной недостаток такого подхода — большая задержка при декодировании и, возможно, чрезмерная сложность реализации декодера. Второй подход связан с использованием каскадных конструкций [10-13]. Схема каскадного блочного кодирования в простом случае состоит в следующем. Сообщения кодируются некоторым корректирующим кодом (внутренним кодом). После этого полученная последовательность кодовых слов внутреннего кода кодируется другим недвоичным кодом (внешним кодом), в котором слова внутреннего кода рассматриваются как недвоичные символы. Основная идея использования каскадирования применительно к многоуровневой памяти состоит в том, чтобы с помощью подходящего, сравнительно простого внутреннего кода обеспечить лучшее согласование внешнего кода с расширенным каналом (включающим кодер, исходный канал и декодер внутреннего кода). Среди возможных вариантов внутреннего кода следует выделить многомерные сигнальные множества, обладающие гибкой структурой с широким диапазоном варьирования параметров и допускающие, как правило, сравнительно простую организацию мягкого декодирования, что может существенно повысить эффективность внешнего кодирования. Весьма важным достоинством каскадной схемы в плане реализации является гибкость ее архитектуры, что позволяет, в частности, реализовать внутреннюю ступень кода, имеющую небольшую сложность, непосредственно на странице кристалла (чипа) флэш-памяти, содержащей защищаемые данные (on-chip implementation).

В работе [14] был предложен и проанализирован ряд каскадных конструкций для флэш-памяти, базирующихся на использовании многомерных решеток в качестве внутреннего кода и кодов Рида — Соломона во внешней ступени каскадной схемы. В качестве основного критерия при выборе конструкций рассматривалась плотность записи, из-

меряемая в количестве бит, приходящихся на одну ячейку памяти. В частности, было показано, что для модели флэш-памяти с четырьмя уровнями записи в ячейке при требуемом уровне вероятности ошибки на бит 10^{-12} предложенные каскадные конструкции обеспечивают плотность записи в диапазоне от 1.93 бит/ячейка (для сравнительно хороших условий записи) до 1.30 бит/ячейка (для наихудших сочетаний числа циклов перезаписи и времени хранения данных). Как показал сравнительный анализ предложенных конструкций, выбор предпочтительного внутреннего кода (и необходимой корректирующей способности внешнего кода Рида — Соломона) существенно зависит от числа циклов перезаписи и времени хранения данных, причем доминирующим фактором является число циклов перезаписи. Так, при времени хранения данных 1 месяц и числе циклов перезаписи, не превосходящем значений порядка 10^4 , предпочтительным является безызбыточный внутренний код (с плотностью записи 2 бит/ячейка и наименьшей помехоустойчивостью); при этом требуемая вероятность ошибки достигается при использовании внешнего кода Рида — Соломона (1024, 988), исправляющего 18 ошибок. При дальнейшем возрастании числа циклов перезаписи, предпочтительными оказываются конструкции с более помехоустойчивым внутренним кодом (вероятность ошибки на $1 \dots 2$ порядка ниже, чем для безызбыточного кода) и, соответственно, менее жесткими требованиями к помехоустойчивости внешнего кода Рида — Соломона — $6 \dots 7$ исправляемых ошибок вместо 18.

Отличительной особенностью конструкций, обеспечивающих максимизацию плотности записи, является сравнительно высокая вероятность ошибки декодирования внутреннего кода, что, в свою очередь, требует достаточно большой корректирующей способности внешнего кода. В результате, как показывает предварительный анализ сложности реализации таких конструкций, сложность каскадного декодера практически полностью определяется сложностью декодирования внешнего кода Рида — Соломона. Так, для внутренних кодов на основе решеток D_5 и E_7 (целочисленных решеток, порождаемых кодом с проверкой на четность (5,4) и кодом Хэмминга (7,4) соответственно) сложность реализации декодера можно примерно оценить (с использованием данных, представленных в [15]), как 1 тысяча и 1.5 тысячи цифровых логических вентилях (4-transistor NAND gates) соответственно. Представление о порядке сложности реализации декодирования кодов Рида — Соломона (512,498) и (2048,2036), используемых в качестве внешних кодов для указанных внутренних кодов, можно получить на основе доступных сведений о реализации декодеров ряда широко используемых кодов Рида — Соломона. Согласно публикациям [16, 17], сложность реализации декодера кода Рида — Соломо-

на (255,239), исправляющего 8 ошибок, примерно оценивается в диапазоне от 39 тысяч до 124 тысяч (в зависимости от степени параллелизации вычислений) цифровых логических вентилях.

Столь высокая сложность декодирования внешнего кода — по меньшей мере несколько десятков тысяч цифровых логических вентилях — может оказаться серьезным ограничением даже при реализации внешней ступени кодирования вне чипа (off-chip). Если в качестве ключевого параметра, определяющего возможность использования помехоустойчивого кодирования для многоуровневой флэш-памяти, выступает сложность декодирования, то в общем классе предложенных в [14] каскадных конструкций следует выделить конструкции с минимальной, по возможности, сложностью декодирования. Как отмечалось выше, сложность каскадного декодера практически полностью определяется сложностью декодирования внешнего кода Рида — Соломона. В свою очередь, сложность декодера кода Рида — Соломона, как известно, степенным образом возрастает с увеличением числа исправляемых ошибок. Это означает, что один из основных путей снижения сложности каскадного декодера — использование в рассматриваемой каскадной схеме кодов Рида — Соломона с небольшим числом исправляемых ошибок. Так, согласно результатам, представленным в [18], сложность реализации декодера кода Рида — Соломона (255,249), исправляющего 3 ошибки, примерно оценивается в 15 тысяч двухвходовых вентилях.

В данной работе рассматривается каскадная схема, в которой внешний код Рида — Соломона исправляет малое число ошибок — не более 4...5. Известно, что при таком ограничении вычисление коэффициентов многочлена локаторов ошибок может быть реализовано с использованием явных и достаточно простых алгебраических выражений, что и подтверждают результаты работы [18]. В случае исправления большего числа ошибок для реализации этого этапа декодирования требуется выполнение более сложных в реализации алгоритма Берлекэмп — Месси или алгоритма Питерсона. Снижение корректирующей способности внешнего кода должно сопровождаться повышением помехоустойчивости внутренней ступени кодирования. Это требование реализуется переходом к более широкому классу внутренних кодов, а именно кодов на основе решеток Барнса — Уолла [19]. Учитывая, что для внутренних кодов на основе решеток Барнса — Уолла сложность декодирования невелика [19], можно ожидать заметного снижения сложности реализации каскадного декодера при относительном возрастании корректирующей способности внутренней ступени каскадного кода.

При получении характеристик помехоустойчивости предложенной каскадной схемы сначала оценивается вероятность ошибки деко-

дирования внутреннего кода, а затем — внешнего. Вероятность ошибки декодирования внешнего кода оценивается с использованием обычной техники суммирования биномиального распределения, а вероятность ошибки декодирования внутреннего кода — с использованием оригинального подхода, разработанного ранее в [14]. Он заключается в комбинации аддитивного неравенства и оценки вероятности ошибки для кода из двух слов, которая построена с использованием характеристических функций распределений, задающих канал. Кроме того, при вычислении аддитивной границы для вероятности ошибки мы в ряде случаев используем параллельную структуру кодовой решетки внутреннего кода, что позволяет понизить сложность вычислений и ускорить получение окончательного результата.

2. Краткое описание модели многоуровневой флэш-памяти.

Как и в работе [14], анализ проводится в рамках модели, в которой блок флэш-памяти рассматривается как множество независимых ячеек. При этом физический носитель задается как стационарный канал без заметной тенденции к пакетированию ошибок, поэтому математическая модель представляет собой модель канала без памяти, которая полностью определяется моделью одиночной ячейки флэш-памяти [20], а также [29] и ссылки там.

Входные уровни каждой ячейки принимают некоторые фиксированные значения x_0, x_1, \dots, x_{q-1} , а выходные значения представляют собой случайные величины. Распределения этих случайных величин параметризованы значениями x_0, x_1, \dots, x_{q-1} , и они описываются условными функциями плотности вероятности (ф.п.в.) $p_{y|x}(y|x)$, $-\infty < y < \infty$, $x = x_0, x_1, \dots, x_{q-1}$. Среди факторов, определяющих распределение значений выходных уровней ячеек, наиболее значимыми являются: а) начальное распределение пороговых значений; б) количество циклов записи/стирания (program/erasure cycling, P/E cycling); в) взаимная интерференция ячеек (cell-to-cell interference); г) длительность времени хранения (retention) [21, 22]. Как отмечалось в работе [14], интегральное влияние этих факторов может быть достаточно адекватно аппроксимировано гауссовой моделью. Поэтому будем полагать, что модель ячейки определяется ф.п.в. $p_{y|x}(y|x_i)$, которые могут быть аппроксимированы гауссовскими плотностями, то есть в принятой модели:

$$p_{y|x}(y|x_i) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{(y-x_i)^2}{2\sigma_i^2}\right).$$

Такое описание также называется моделью с гауссовским шумом, стандартное отклонение которого σ_i зависит от входного значения x_i (input-dependent additive Gaussian noise, ID-AGN). Математические ожидания x_i и стандартные отклонения σ_i , $i = 0, 1, \dots, q-1$, условных ф.п.в. $p_{y|x}(y|x_i)$ представляют собой основные параметры этой модели. Важной особенностью модели ячейки многоуровневой флэш-памяти является то, что с ростом числа циклов перезаписи N и времени хранения T значения x_i (кроме x_0) уменьшаются, а значения σ_i (кроме σ_0) увеличиваются, что соответствует ухудшению канала с ростом величин T и N . Дальнейшие детали описания модели представлены в [14, 29] и ссылки там. Распространенным примером, используемым далее в качестве основного, служит значение $q = 4$.

3. Схема каскадного кодирования. В рассматриваемой схеме каскадного кодирования используются две ступени: внутренняя и внешняя. Внешний код представляет собой расширенный (удлиненный на один символ) код Рида — Соломона, а внутренний код построен как множество последовательностей, символами которых являются элементы множества $\{0, 1, \dots, q-1\}$, где q — число уровней записи ячейки флэш-памяти. Кодовые символы внешнего кода подвергаются блоковому перемежению, а именно: h последовательных слов кода Рида — Соломона записываются в прямоугольную таблицу, содержащую h строк, после чего последовательно считываются по столбцам. Каждый столбец этой таблицы отображается в один символ внутреннего кода. Параметр h определяется объемом внутреннего кода и размером алфавита кода Рида — Соломона. При $h = 1$ один символ внешнего кода отображается в одно слово внутреннего кода, то есть имеет место обычная каскадная схема кодирования.

Слова внутреннего кода (т.е. символы внешнего кода) строятся как точки *конечного подмножества* B_0 некоторой исходной n -мерной *целочисленной решетки* (lattice) Λ_0 , $B_0 \subset \Lambda_0$.

Замечание. Термину «решетка» соответствуют два различных понятия, используемые в теории связи и теории кодирования, которые в английском переводе записываются как «lattice» и «trellis». В данном тексте используются оба эти понятия, поэтому для устранения неопределенности наряду с русским словом «решетка» местами будет указываться его английский эквивалент, соответствующий контексту.

Объем алфавита и длина слова внутреннего кода равны в этом случае числу уровней записи ячейки флэш-памяти q и размерности

решетки (lattice) n соответственно. Если обозначить объем внутреннего кода (число точек в конечном подмножестве целочисленной решетки $B_0, B_0 \subset \Lambda_0$) как $|B_0|$, то скорость внутреннего кода может быть записана как $r = \log_2 |B_0| / n$. Поскольку $|B_0| \leq q^n$, то имеет место граница $r \leq \log_2 q / n$. Если обозначить через R_1 скорость внешнего кода, то общая скорость каскадного кода, определяющая значение *плотности записи* в ячейку флэш-памяти, может быть вычислена $R_{\text{общ}} = rR_1 = R_1 \log_2 |B_0| / n$, бит/ячейка.

Пусть длина внешнего кода обозначена как N_1 . Предполагая, что внешний код представляет собой расширенный код Рида — Соломона, можно записать, что $N_1^n = |B_0|$.

4. Внутренний код на основе решеток Барнса — Уолла. Обозначим через A множество, состоящее из q целых чисел $A = \{0, 1, \dots, q-1\}$, и будем полагать, что значение $m = \log_2 q$ целое. Множество A , очевидно, соответствует совокупности равномерно распределенных уровней q -уровневой ячейки памяти. Множество A связано с реальным множеством целевых уровней q -уровневой ячейки памяти $X = \{x_0, x_1, \dots, x_{q-1}\}$ посредством взаимно однозначного отображения I множества A на множество X вида $I(i) = x_i$, $0 \leq i \leq q-1$. Пусть B есть прямое произведение n экземпляров множества A , $B = A^n$. Множество B будем называть сигнальным множеством рассматриваемой далее каскадной конструкции. Общее число точек в сигнальном множестве B равно $\exp_2(mn)$, так что множество B соответствует n ячейкам памяти, хранящим $m \times n$ бит. Внутренний код B_0 рассматриваемой далее каскадной кодовой конструкции представляет собой некоторое подмножество сигнального множества B , $B_0 \subseteq B$. При построении и анализе внутреннего кода удобно рассматривать исходное сигнальное множество B как конечное подмножество n -мерной целочисленной решетки Z^n , а внутренний код B_0 — как конечное подмножество решетки $\Lambda_0 \subseteq Z^n$. Такой подход позволяет получить регулярную процедуру построения внутренних кодов на основе разбиений исходной решетки Λ_0 на подходящие вложенные подрешетки. Подрешетка Λ_0 определяет разбиение Z^n / Λ_0 решетки Z^n на подрешетку Λ_0 и ее смежные классы; пусть порядок этого разбиения

ния $|Z^n / \Lambda_0| = \exp_2(r_0)$. Внутренний код B_0 есть пересечение решетки Λ_0 и сигнального множества B , $B_0 = \Lambda_0 \cap B$. Очевидно, общее число точек в множестве B_0 равно $\exp_2(mn - r_0)$.

Будем полагать, что решетка Λ_0 есть: 1) mod-4 решетка, то есть $\Lambda_0 \supseteq 4Z^n$; и 2) является разложимой решеткой. Разложимость mod-4 решетки Λ_0 означает, что Λ_0 может быть представлена в следующем виде [19]:

$$\Lambda_0 = 4Z^n + 2\{\mathbf{a}_l^{(1)} \mathbf{G}_1 \mid 1 \leq l \leq \exp_2(k_1)\} + \{\mathbf{a}_l^{(0)} \mathbf{G}_0 \mid 1 \leq l \leq \exp_2(k_0)\}, \quad (1)$$

где $\mathbf{a}_l^{(1)}$ и $\mathbf{a}_l^{(0)}$ — двоичные (над алфавитом $\{0,1\}$) векторы размерности k_1 и k_0 соответственно; \mathbf{G}_1 и \mathbf{G}_0 — двоичные $k_1 \times n$ и $k_0 \times n$ матрицы соответственно, причем \mathbf{G}_0 есть подматрица матрицы \mathbf{G}_1 ; и $k_1 + k_0 = 2n - r_0$. Матрицы \mathbf{G}_1 и \mathbf{G}_0 в разложении (1) можно интерпретировать как порождающие матрицы некоторых линейных кодов C_1 и C_0 соответственно, поэтому схематично разложение (1) можно представить в виде:

$$\Lambda_0 = 4Z^n + 2C_1 + C_0, \quad (2)$$

где код C_0 есть подкод кода C_1 . Квадрат минимального евклидова расстояния между точками решетки Λ_0 может быть найден как $\Delta_0^2 = \min\{16, 4d_H(C_1), d_H(C_0)\}$. В частном случае, когда код C_1 безызбыточный, схематичное представление разложения (1) приобретает вид $\Lambda_0 = 2Z^n + C_0$. При этом квадрат минимального евклидова расстояния между точками решетки Λ_0 равен $\Delta_0^2 = \min\{4, d_H(C_0)\}$.

Приведем два примера использования представления (2). Например, если в качестве Λ_0 выбрана решетка RE_8 , полученная путем прореживания решетки Gosset E_8 , код C_1 в (2) есть код проверки на четность (8,7,2), а код C_0 представляет собой код с повторением (8,1,8). Второй пример: Λ_0 есть 16-мерная решетка Барнса — Уолла Λ_{16} ; тогда код C_1 в выражении (2) есть код проверки на четность (16,15,2), а код C_0 — код Рида — Маллера (16,5,8). В обоих

примерах квадрат минимального евклидова расстояния $\Delta_0^2 = 8$. Заметим, что для рассматриваемой модели многоуровневой флэш-памяти [14] евклидовы расстояния лишь косвенным образом определяют вероятность ошибки декодирования внутреннего кода. Тем не менее эти характеристики оказываются полезными при конструировании подходящих каскадных схем кодирования.

Разложение (1) решетки Λ_0 однозначно определяет аналогичное разложение внутреннего кода B_0 . Пусть $B_1 = 4Z^n \cap B$, то есть:

$$B_1 = \{0, 4, \dots, q - 4\}^n; \quad (3)$$

в частности, при $q = 4$ множество B_1 состоит из одного нулевого вектора. Тогда внутренний код B_0 может быть представлен в следующем виде:

$$B_0 = B_1 + 2\{\mathbf{a}_l^{(1)} \mathbf{G}_1 \mid 1 \leq l \leq \exp_2(k_1)\} + \{\mathbf{a}_l^{(0)} \mathbf{G}_0 \mid 1 \leq l \leq \exp_2(k_0)\}, \quad (4)$$

где, как и для представления (1), $\mathbf{a}_l^{(1)}$ и $\mathbf{a}_l^{(0)}$ — двоичные (над алфавитом $\{0,1\}$) векторы размерности k_1 и k_0 соответственно; \mathbf{G}_1 и \mathbf{G}_0 — двоичные $k_1 \times n$ и $k_0 \times n$ матрицы соответственно; и $k_1 + k_0 = 2n - r_0$. Матрицы \mathbf{G}_1 и \mathbf{G}_0 , как и в разложении (1), есть порождающие матрицы линейных кодов C_1 и C_0 соответственно, поэтому схематично разложение (4) можно представить в виде:

$$B_0 = B_1 + 2C_1 + C_0. \quad (5)$$

Представления (4) и (5), позволяют явно указать на весьма простую процедуру кодирования для внутреннего кода B_0 . Двоичный блок \mathbf{u} , кодируемый кодом B_0 , состоит, очевидно, из $mn - r_0$ символов. Представим блок \mathbf{u} в виде набора m подблоков:

$$\mathbf{u} = (\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_{m-1}), \quad (6)$$

где подблок \mathbf{u}_0 имеет длину k_0 , подблок \mathbf{u}_1 имеет длину k_1 , $k_1 + k_0 = 2n - r_0$ остальные подблоки $\mathbf{u}_2, \dots, \mathbf{u}_{m-1}$ имеют длину n . Положим, что кодовое слово \mathbf{b} внутреннего кода B_0 , соответствующее входному кодируемому блоку \mathbf{u} , имеет вид:

$$\mathbf{b} = \mathbf{u}_0 \mathbf{G}_0 + 2\mathbf{u}_1 \mathbf{G}_1 + \sum_{k=2}^{m-1} 2^k \mathbf{u}_k, \quad (7)$$

где \mathbf{G}_0 — порождающая матрица линейного двоичного (n, k_0) кода C_0 , \mathbf{G}_1 — порождающая матрица линейного двоичного (n, k_1) кода C_1 в (5). Нетрудно видеть, что совокупность из $\exp_2(k_1 + k_0 + (m-2)n) = \exp_2(mn - r_0)$ n -мерных векторов над алфавитом A , порождаемых согласно (7), совпадает с множеством слов внутреннего кода B_0 , задаваемого представлением (4). Скорость внутреннего кода $r = (k_0 + k_1) / n + m - 2$, бит/ячейка. Квадрат минимального евклидова расстояния внутреннего кода δ^2 совпадает, очевидно, с Δ_0^2 и равен:

$$\delta_0^2 = \min \{16, 4d_H(C_1), d_H(C_0)\}.$$

Внутренний код B_0 , как следует из (3) и соотношения (7), можно рассматривать как прямую сумму кодовых слов линейного двоичного кода C_0 (с порождающей матрицей \mathbf{G}_0), кодовых слов линейного двоичного кода C_1 (с порождающей матрицей \mathbf{G}_1), умноженных на 2, и подмножества (3). Это означает, что при описании внутреннего кода в виде *обычной* решетчатой диаграммы Ω , состоящей из n ярусов (bit-level trellis diagram), диаграмма Ω может быть представлена в виде прямого произведения обычных решетчатых диаграмм [23] кодов C_0 и C_1 , а также «кода» B_1 :

$$\Omega = \Omega_0 \otimes \Omega_1 \otimes \Omega_2,$$

где $\Omega_0, \Omega_1, \Omega_2$ — решетчатые диаграммы кодов C_0 , C_1 и «кода» B_1 соответственно. Учитывая, что «код» B_1 определен выражением (3), диаграмма Ω может быть представлена в виде произведения $\Omega_0 \otimes \Omega_1$ решетчатых диаграмм кодов C_0 и C_1 с $m-1$ параллельными ребрами для каждой пары смежных состояний. Ребра диаграммы Ω внутреннего кода B_0 маркируются символами из множества A ; при этом, как следует из равенства (7), параллельные ребра, соединяющие пару смежных состояний, маркируются символами одного из следующих четырех множеств: $\{0, 4, \dots, q-4\}$, $\{1, 5, \dots, q-3\}$, $\{2, 6, \dots, q-2\}$, $\{3, 7, \dots, q-1\}$.

Представление кода с использованием обычной (bit-level) решетчатой диаграммы является, как хорошо известно из [23], одним из ряда возможных. Во многих случаях более предпочтительным может оказаться описание кода B_0 посредством секционированной решетчатой диаграммы, при котором диаграмма кода формируется из s секций. Каждая из этих секций объединяет n/s ярусов исходной bit-level диаграммы, при этом количество параллельных ребер, соединяющих каждую пару смежных состояний, составляет $(m-1)^{n/s}$. Отметим также, что в ряде случаев прямое произведение решетчатых диаграмм кодов C_0 и C_1 реализуется в виде параллельных структурно идентичных подрешеток без перекрестных связей между ними.

При введении внешнего кодирования блок \mathbf{u} (6), кодируемый внутренним кодом, есть двоичное представление одного или нескольких символов внешнего кода. Будем полагать, что внутренний код B_0 с минимальным евклидовым расстоянием Δ_0 используется для кодирования h символов перемеженных кодовых слов (по одному из каждого слова) внешнего расширенного кода Рида — Соломона. В этом случае, блок \mathbf{u} , кодируемый кодом B_0 , есть двоичное представление этих h символов, принадлежащих h словам внешнего кода Рида — Соломона. Учитывая, что блок \mathbf{u} состоит из $mn-r_0$ символов, отношение $(mn-r_0)/h$ должно быть целым, а максимально возможная длина внешнего кода Рида — Соломона составляет $\exp_2((mn-r_0)/h)$. Отображение блока \mathbf{u} , представляющего собой в двоичной форме h символов перемеженных кодовых слов (по одному из каждого слова) в кодовое слово внутреннего кода \mathbf{b} производится в соответствии с правилом (7). Плотность записи при таком варианте внешнего кодирования составляет $R_1(m-r_0/n)$ бит/ячейка, где R_1 — скорость кода Рида — Соломона. В таблице 1 приведены параметры ряда каскадных конструкций для $q=4$ с внутренним кодом на основе следующих решеток (или подрешеток) Барнса — Уолла [19]: $BW_8(0,2) = E_8$, $BW_8(-1,2) = RE_8$, $BW_{16}(1,3) = H_{16}$, $BW_{16}(0,3) = \Lambda_{16}$, $BW_{16}(-1,3) = RH_{16}$, $BW_{32}(1,4) = H_{32}$, $BW_{32}(0,4) = \Lambda_{32}$; оператор R в обозначениях подрешеток $BW_8(-1,2)$ и $BW_{16}(-1,3)$ представляет собой оператор вращения, определенный в [10]. Компонентные коды C_0 и C_1 в этом случае представляют собой коды Рида — Маллера [19].

Отметим, что все внутренние коды, представленные в таблице 1, могут быть описаны посредством секционированной решетчатой диаграммы с четырьмя секциями, каждая из которых объединяет, соответственно, 2, 4 и 8 ярусов исходной bit-level диаграммы.

Таблица 1. Параметры каскадных конструкций с внутренним кодом на основе решеток (или подрешеток) Барнса-Уолла

Исходная решетка Λ_0 и ее размерность n	Коды C_0 и C_1	Мин. расстояние δ	Объем внутреннего кода $ B_0 $	Возможные значения h и длин кода Рунда-Соломона N_1	Плотность записи, бит/ячейка
$E_8, n=8$	$C_0=(8,4,4)$ $C_1=(8,8,1)$	2	2^{12}	$N_1=16, h=3$ $N_1=64, h=2$	$1.5R_1$
$RE_8, n=8$	$C_0=(8,1,8)$ $C_1=(8,7,2)$	$\sqrt{8}$	2^8	$N_1=256, h=1$	$1.0R_1$
$H_{16}, n=16$	$C_0=(16,11,4)$ $C_1=(16,16,1)$	2	2^{27}	$N_1=512, h=3$	$1.688R_1$
$\Lambda_{16}, n=16$	$C_0=(16,5,8)$ $C_1=(16,15,2)$	$\sqrt{8}$	2^{20}	$N_1=16, h=5$ $N_1=32, h=4$ $N_1=1024, h=2$	$1.25R_1$
$R\Lambda_{16}, n=16$	$C_0=(16,1,16)$ $C_1=(16,11,4)$	4	2^{12}	$N_1=64, h=2$ $N_1=4096, h=1$	$0.75R_1$
$H_{32}, n=32$	$C_0=(32,16,8)$ $C_1=(32,31,2)$	$\sqrt{8}$	2^{47}	$N_1=2^{47}, h=1$	$1.468R_1$
$\Lambda_{32}, n=32$	$C_0=(32,6,16)$ $C_1=(32,26,4)$	4	2^{32}	$N_1=16, h=8;$ $N_1=256, h=4$	$1.0R_1$

Наибольший интерес представляют коды, обеспечивающие, с одной стороны, высокую плотность записи, а с другой — большое значение минимального расстояния. Среди них далее мы рассматриваем:

а) код на основе решетки E_8 , для которого $\delta=2$ и плотность записи $1.5R_1$ бит/ячейка;

б) код на основе решетки RE_8 , для которого $\delta=\sqrt{8}$ и плотность записи $1.0R_1$ бит/ячейка;

с) код на основе решетки Λ_{16} , для которого $\delta=\sqrt{8}$ и плотность записи $1.25R_1$ бит/ячейка.

Сложность реализации кода на основе решетки H_{32} , для которого $\delta=\sqrt{8}$ и плотность записи $1.468R_1$ бит/ячейка, представляется слишком большой, и поэтому он исключается из дальнейшего рассмотрения. Рассмотрим подробно структуру решетчатых диаграмм кодов «а», «б» и «с».

Код «а». Этот код строится на основе минимальной решетчатой диаграммы двоичного кода Хэмминга (8,4,4), показанной на рисунке 1[24, 25]. Для получения из нее решетчатой диаграммы внутреннего кода ребра, отмеченные символом 0 (показаны на рисунке пунктиром), заменяются двумя параллельными ребрами, которые отмечаются символами 0 и 2, а ребра, отмеченные символом 1 (показаны на рисунке сплошной линией), — двумя параллельными ребрами, отмеченными символами 1 и 3.

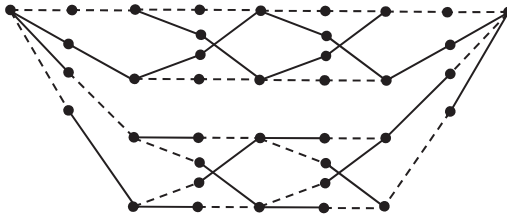


Рис. 1. Решетчатая диаграмма двоичного кода Хэмминга (8,4,4)

Код «b». Решетчатая диаграмма этого кода строится на основе решетчатых диаграмм кода с одной проверкой на четность (8,7,2) и кода с повторением (8,1,8). Иллюстрация показана на рисунке 2.

Код «с». Диаграмма этого кода строится на основе решетчатой диаграммы кода Рида — Маллера (16,5,8). Она в секционированном виде приведена на рисунке 3 [19]. Эта диаграмма, как видно из рисунка, представляет собой четыре параллельных решетки меньшей сложности без перекрестных связей между ними.

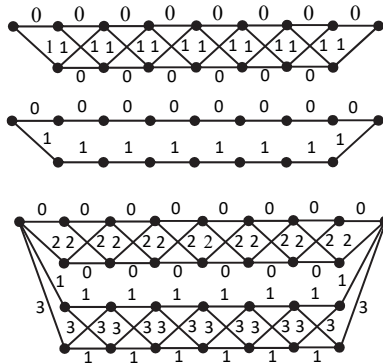


Рис. 2. Решетчатая диаграмма двоичного кода (8,7,2), двоичного кода (8,1,8) и внутреннего кода «b»

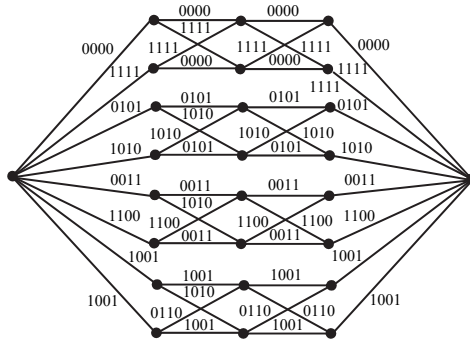


Рис. 3. Решетчатая диаграмма двоичного кода (16,5,8)

Структура решетки меньшей сложности (малой решетки) в секционированном виде и в виде обычной диаграммы показана на рисунке 4. Вектор \mathbf{a} принимает значения (0000), (0101), (0011) и (1001), и в зависимости от его значения получаются четыре варианта малой решетки, из которых строится решетка, показанная на рисунке 3. В нижней части рисунка 4, где показана малая решетка в виде обычной или несекционированной диаграммы, разметка ребер не указана из-за недостатка места. Эта разметка строится очевидным образом и может быть легко восстановлена.

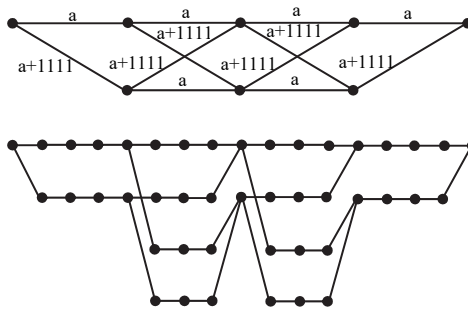


Рис. 4. Малая решетка в секционированном виде и обычном представлении, $\mathbf{a} = (0000), (0101), (0011), (1001)$

Полная решетчатая диаграмма кода «в» строится как параллельная структура решеток, полученных в результате комбинации диаграмм кода с проверкой на четность (16,15,2) и малой решетки, показанной на рисунке 5. Эта диаграмма изображена на рисунке 6.

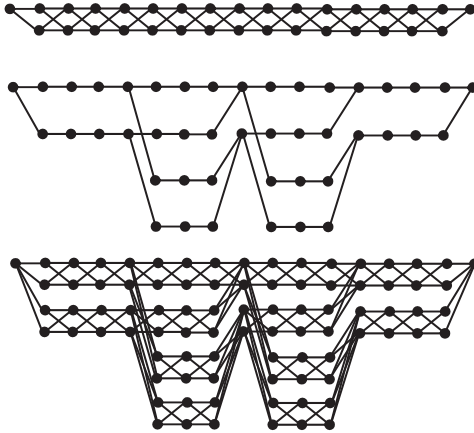


Рис. 5. Решетчатая диаграмма двоичного кода $(16,15,2)$, малой решетки двоичного кода $(16,5,8)$, и их комбинация

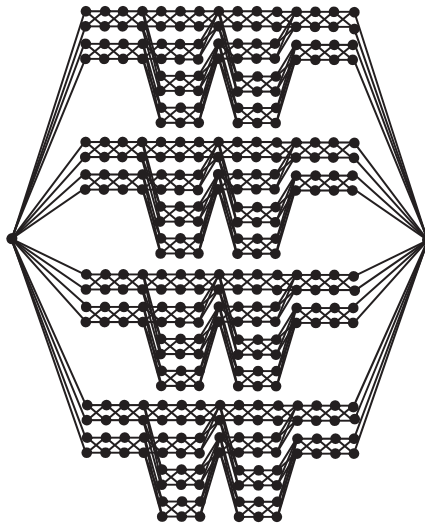


Рис. 6. Решетчатая диаграмма кода «с»

Замечание. Варианты, представленные в таблице 1, фактически исчерпывают перечень приемлемых с точки зрения приложений кон-

струкций внутреннего кода на основе решеток Барнса — Уолла. Расширение перечня приемлемых конструкций внутреннего кода может быть получено по меньшей мере двумя путями. Первый — рассмотрение произвольных mod-4 решеток, порождаемых (в соответствии с (1) и (2)) линейными двоичными кодами C_0 и C_1 , такими, что код C_0 есть подкод кода C_1 . Второй путь — конструирование внутреннего кода на основе правила (7) с использованием подходящих (для рассматриваемой ситуации) кодов C_0 и C_1 . При этом множество Θ , порождаемое как:

$$\Theta = 4Z^n + 2\{\mathbf{a}_l^{(1)}\mathbf{G}_1 \mid 1 \leq l \leq \exp_2(k_1)\} + \{\mathbf{a}_l^{(0)}\mathbf{G}_0 \mid 1 \leq l \leq \exp_2(k_0)\},$$

где $\mathbf{a}_l^{(1)}$ и $\mathbf{a}_l^{(0)}$ — двоичные (над алфавитом $\{0, 1\}$) векторы размера k_1 и k_0 соответственно; \mathbf{G}_1 — порождающая $k_1 \times n$ матрица кода C_1 , \mathbf{G}_0 — порождающая $k_0 \times n$ матрица кода C_0 , в общем случае не является решеткой, то есть аддитивной группой. Кроме того, канал записи/считывания для флэш-памяти представляет собой, как отмечалось в [14], ID-AGN канал, поэтому вычисление (или оценивание) вероятности ошибки для кода B_0 , описываемого решетчатой диаграммой Ω , требует оперирования с диаграммой $\Omega \otimes \Omega$.

5. Введение внешнего кодирования. Кодовые символы внешнего кода Рида — Соломона подвергаются, как указывалось выше, блоковому перемежению таким образом, что внутренний код B_0 используется для кодирования h символов перемеженных кодовых слов (по одному из каждого слова); параметр h определяется объемом внутреннего кода и размером алфавита (длиной N_1) расширенного кода Рида — Соломона: $N_1^h = |B_0|$. В результате исключается пакетирование ошибок на входе декодера кода Рида — Соломона, то есть ошибки в кодовых символах кода Рида — Соломона являются независимыми, несмотря на то, что слово внутреннего кода есть результат кодирования h , $h \geq 1$ символов кода Рида — Соломона. Среди представленных в таблице 1 каскадных конструкций для дальнейшего детального анализа были выбраны три конструкции (с четырьмя вариантами для каждой), представленные в таблице 2. Этот выбор был сделан, исходя из соотношения между минимальным расстоянием, плотностью записи и ограничением на число управляемых внешним кодом ошибок, $t \leq 4$.

Таблица 2. Параметры выбранных для анализа каскадных конструкций

Вариант внутреннего кода	Исходная решётка Λ_0 и ее размерность n	Мин. расстояние δ	Объем внутреннего кода $ B_0 $	Коды Рида-Соломона	Плотность записи, бит/ячейка
«а»	$E_8, n=8$	$\sqrt{2}$	2^{12}	$(64,62), h=2$ $(64,60), h=2$ $(64,58), h=2$ $(64,56), h=2$	1.453 1.406 1.359 1.312
«б»	$RE_8, n=8$	$\sqrt{8}$	2^8	$(256,254), h=1$ $(256,252), h=1$ $(256,250), h=1$ $(256,248), h=1$	0.992 0.984 0.977 0.969
«с»	$\Lambda_{16}, n=16$	$\sqrt{8}$	2^{20}	$(1024,1022), h=2$ $(1024,1020), h=2$ $(1024,1018), h=2$ $(1024,1016), h=2$	1.248 1.245 1.243 1.240

6. Анализ помехоустойчивости. Анализ помехоустойчивости предложенных каскадных кодовых конструкций включает в себя два этапа. Первый этап состоит в вычислении (оценке) вероятности ошибки декодирования по максимуму правдоподобия (МП) слова внутреннего кода. На втором этапе вычисляется (оценивается) вероятность ошибки декодирования слова внешнего кода с использованием результатов, полученных на первом этапе. Рассмотрим оба эти этапа последовательно.

Вычисление аддитивной границы ошибки декодирования внутреннего кода, представленного параллельными решетками. Некоторые из отобранных ранее вариантов внутреннего кода задаются с использованием решетчатой диаграммы, имеющей вид параллельных решетчатых структур, не связанных друг с другом. Эта особенность их структуры может быть использована для уменьшения сложности вычисления верхней границы вероятности ошибки декодирования внутреннего кода.

Аддитивная граница вероятности ошибки декодирования по МП слова внутреннего кода задается выражением [14]:

$$p_e \leq \frac{1}{\pi M} \int_0^\infty \operatorname{Re} \frac{D(\alpha - j\beta) - M}{\beta + j\alpha} d\alpha, \quad (8)$$

где

$$D(\omega) = \sum_x \sum_{x'} \prod_{l=1}^n c_{z(x^{(l)}, x'^{(l)})}(\omega). \quad (9)$$

M — количество слов внутреннего кода, n — длина слова внутреннего кода. Суммирование в правой части равенства (9) в обеих суммах выполняется по всем парам слов внутреннего кода \mathbf{x} и \mathbf{x}' , $\mathbf{x} = (x^{(1)}, \dots, x^{(n)})$, $\mathbf{x}' = (x'^{(1)}, \dots, x'^{(n)})$. Функции $c_{z(x^{(l)}, x'^{(l)})}(\omega)$ в (9) представляют собой характеристические функции величин $z(y|x, x') = \ln(p_{y|x}(y|x') / p_{y|x}(y|x))$, то есть:

$$c_{z(y|x, x')}(\omega) = e^{\overline{j\omega z(y|x, x')}},$$

где усреднение выполняется по значениям выходной величины канала y , то есть по значению, считанному из ячейки флэш-памяти. При условии, что было передано (записано) значение уровня x , считанное значение y имеет гауссовское распределение с математическим ожиданием x и стандартным отклонением $\sigma(x)$. Заметим, что $\sigma(x) = \sigma_i$, если $x = x_i$. Показано [14], что при $\sigma(x) \neq \sigma(x')$ функция $c_{z(x, x')}(\omega)$ задается следующим равенством:

$$c_{z(y|x, x')}(\omega) = \frac{\sigma(x)^{j\omega} \sigma(x')^{1-j\omega}}{\sqrt{j\omega\sigma^2(x) + (1-j\omega)\sigma^2(x')}} \times \exp\left(-\frac{j\omega(1-j\omega)(x-x')^2}{2(j\omega\sigma^2(x) + (1-j\omega)\sigma^2(x'))}\right), \quad (10)$$

при условиях:

$$\text{Im}\omega < \sigma^2(x') / (\sigma^2(x) - \sigma^2(x')), \text{ если } \sigma(x) > \sigma(x')$$

и

$$-\text{Im}\omega < \sigma^2(x') / (\sigma^2(x') - \sigma^2(x)), \text{ если } \sigma(x) < \sigma(x').$$

В случае, когда $\sigma(x) = \sigma(x')$, характеристическая функция $c_{z(x, x')}(\omega)$ вычисляется согласно выражению:

$$c_{z(y|x, x')}(\omega) = \exp\left(-\frac{j\omega(1-j\omega)(x-x')^2}{2\sigma^2(x)}\right).$$

Обозначим для краткости записи произведение в правой части (9) как:

$$C_{Z(y|x,x')}(\omega) = \prod_{l=1}^n c_{z(y^{(l)}|x^{(l)},x'^{(l)})}(\omega), \quad (11)$$

где случайная величина $Z(y|x,x')$ определена следующим образом:

$$Z(y|x,x') = \sum_{l=1}^n z(y^{(l)}|x^{(l)},x'^{(l)}).$$

Заметим также, что в силу независимости величин $y^{(l)}$ при фиксированных значениях $x^{(l)}, x'^{(l)}$:

$$C_{Z(y|x,x')}(\omega) = e^{\overline{j\omega Z(y|x,x')}}}, \quad (12)$$

где усреднение выполняется по значениям вектора $y = (y^{(1)}, y^{(2)}, \dots, y^{(n)})$. Тогда выражение (9) можно переписать в виде:

$$D(\omega) = \sum_x \sum_{x'} C_{Z(y|x,x')}(\omega). \quad (13)$$

Вычисление величин (13) выполняется по алгоритму, разработанному в [14]. Вкратце он может быть описан следующим образом. Рассмотрим решетку (trellis) внутреннего кода (см. рисунки 1, 2, 6). Обозначим $v(l)$ число вершин на l -ом уровне решетки, $l = 0, 1, \dots, n$. Очевидно, что для всех решеток $v(0) = v(n) = 1$. Например, для решетки, показанной на рисунке 1, $v(0) = v(8) = 1$, $v(1) = v(2) = v(4) = v(6) = v(7) = 4$, $v(3) = v(5) = 8$. Пусть $\chi_l(s, s')$ — множество меток ребер, соответствующих переходам из вершины s в вершину s' на l -ом уровне, $s = 0, 1, \dots, v(l-1) - 1$, $s' = 0, 1, \dots, v(l) - 1$. Полагая, что номера состояний назначаются сверху вниз по порядку, для решетки, показанной на рисунке 1, имеем, например, $\chi_2(1, 1) = \{x_1, x_3\}$, $\chi_2(3, 1) = \emptyset$, а $\chi_2(2, 5) = \{x_0, x_2\}$ и так далее. Определим величины:

$$P_{ij}^{(l)}(\omega) = \sum_{x \in \chi_l(s_a, s_a')} \sum_{x' \in \chi_l(s_b, s_b')} c_{z(x, x')}(\omega),$$

где $i = v(l-1)s_a + s_b$ и $j = v(l)s_a' + s_b'$. Тогда можно записать, что:

$$D(\omega) = \mathbf{a}(\omega)\mathbf{P}_2(\omega)\dots\mathbf{P}_{n-1}(\omega)\mathbf{b}(\omega),$$

где $\mathbf{P}_i(\omega) = [P_{ij}^{(l)}(\omega)]$ — матрица размерности $v(l-1)^2 \times v(l)^2$, $i = 0, 1, \dots, v(l-1)^2 - 1$, $j = 0, 1, \dots, v(l)^2 - 1$, $\mathbf{a}(\omega) = [P_{00}^{(1)}(\omega), \dots, P_{0, v(l)^2 - 1}^{(1)}(\omega)]$, $\mathbf{b}(\omega) = [P_{00}^{(n)}(\omega), \dots, P_{v(n-1)^2 - 1, 0}^{(n)}(\omega)]^T$. Полученные таким образом величины $D(\omega)$ подставляются в выражение (8), и после численного интегрирования получаем окончательную оценку для вероятности ошибки декодирования внутреннего кода.

Выражение (8) справедливо для любых значений параметра β , удовлетворяющих условиям, следующим из определения (10) характеристической функции $c_{z(x,x)}(\omega)$. При практических вычислениях его следует выбирать таким образом, чтобы подынтегральное выражение представляло собой функцию, удобную для численного интегрирования. В частности, значение $\beta = 1/2$ оказывается подходящим для получения вполне удовлетворительных результатов.

В ряде случаев структура внутреннего кода задается решеткой, которая может быть разделена на несколько параллельных малых решеток (см. рисунки 2 и 6). Это обстоятельство можно использовать для упрощения вычисления границы вероятности ошибки (8). В равенстве (13) суммирование выполняется по всем \mathbf{x} и \mathbf{x}' , которые соответствуют решетчатой диаграмме, схематично показанной на рисунке 7.

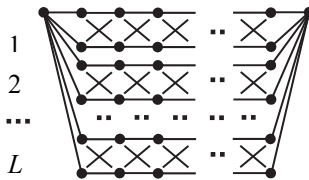


Рис. 7. Решетчатая диаграмма, состоящая из L параллельных решетчатых диаграмм (малых решеток)

С учетом параллельной структуры решетчатой диаграммы выражение (13) можно записать как:

$$D(\omega) = \sum_{i=1}^L \sum_{k=1}^L \sum_{\mathbf{x} \in X_i} \sum_{\mathbf{x}' \in X_k} C_{Z(\mathbf{y}|\mathbf{x}, \mathbf{x}')}(\omega) = \sum_{i=1}^L \sum_{k=1}^L D_{ik}(\omega), \quad (14)$$

где X_i — множество слов внутреннего кода, соответствующих i -ой малой решетке, $i = 1, 2, \dots, L$ и

$$D_{ik}(\omega) = \sum_{x \in X_i} \sum_{x' \in X_k} C_{Z(y|x, x')}(\omega). \quad (15)$$

Опуская промежуточные выкладки, приведем окончательное выражение для вероятности ошибки декодирования по МП слова внутреннего кода, в котором используются функции (15):

$$p_e \leq \frac{1}{\pi M} \int_0^\infty \operatorname{Re} \frac{\sum_{i=1}^L D_{ii}(\alpha - j/2) + 2 \sum_{i=1}^L \sum_{k=i+1}^L D_{ik}(\alpha - j/2) - M}{1/2 + j\alpha} d\alpha. \quad (16)$$

Отметим, что для вычисления правой части (16) нужно вычислить $L(L+1)/2$ функций $D_{ik}(\cdot)$, а не L^2 , как это следует из формулы (14). Вычисление сумм характеристических функций (15) выполняется по алгоритму, разработанному в [14]. Далее они используются в выражении (16), которое дает окончательную оценку для вероятности ошибки декодирования внутреннего кода, представленного комбинацией параллельных решеток.

На рисунках 8-10 показаны зависимости вероятности ошибки декодирования символа внутреннего кода p_e от числа циклов перезаписи N . Напомним, что зависимость параметров ф.п.в. $p_{y|x}(y|x)$ от числа циклов перезаписи N и времени хранения T задаются выражениями, следующими из описания принятой модели канала записи (см. [14, 29] и ссылки там).

Для указанных примеров скорость (плотность записи) составляет 1.5 бит/ячейка для внутреннего кода «а», построенного на основе решетки E_8 ; 1.0 бит/ячейка для внутреннего кода «б», построенного на основе решетки RE_8 ; и 1.25 бит/ячейка для внутреннего кода «с», построенного на основе решетки Λ_{16} . Из представленных данных следует, что приемлемая вероятность ошибки не может быть достигнута при использовании только внутреннего кодирования, но можно надеяться, что требуемого уровня надежности можно достичь, используя внешнее кодирование с исправлением небольшого числа ошибок.

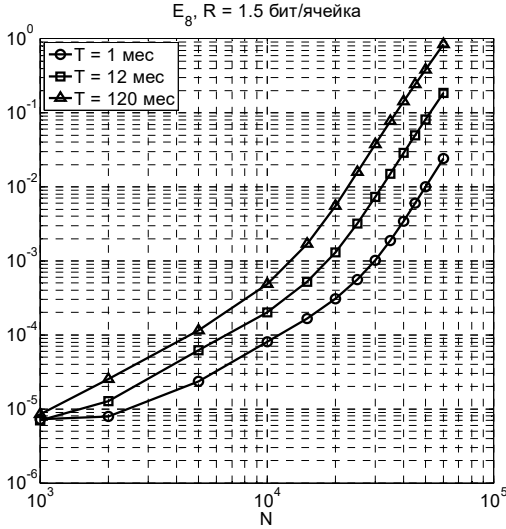


Рис. 8. Вероятность ошибки декодирования символа внутреннего кода «а», построенного на основе решетки E_8 , в зависимости от числа циклов перезаписи

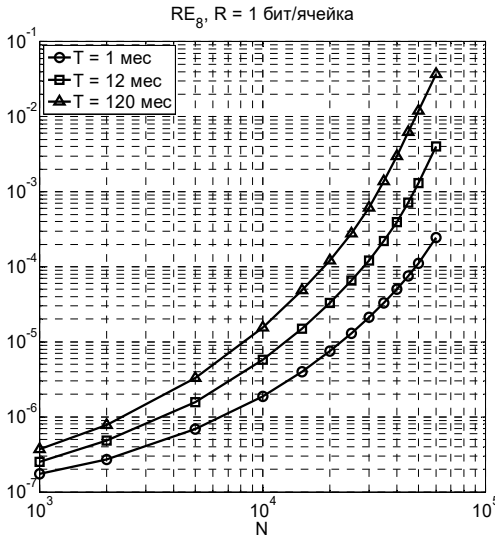


Рис. 9. Вероятность ошибки декодирования символа внутреннего кода «б», построенного на основе решетки RE_8 , в зависимости от числа циклов перезаписи

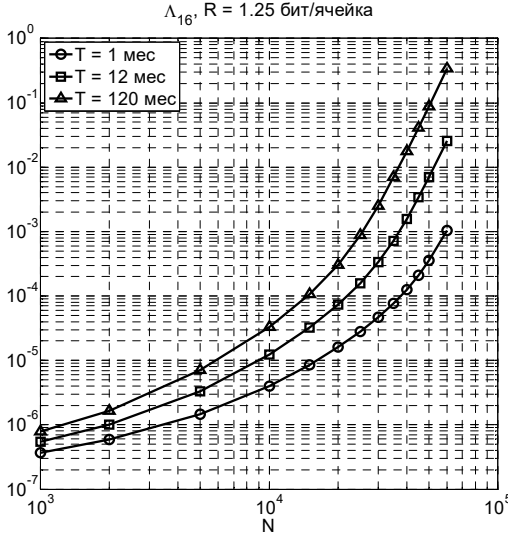


Рис. 10. Вероятность ошибки декодирования символа внутреннего кода «с», построенного на основе решетки Λ_{16} , в зависимости от числа циклов перезаписи

Вероятность ошибки декодирования внешнего кода. Численные результаты. Рассмотрим каскадную конструкцию с внешним кодом, в качестве которого используется расширенный код Рида — Соломона. Примеры таких конструкций перечислены в таблице 2. Оценим вероятность ошибки получения на выходе декодера ошибочного бита (вероятность ошибки на бит). Вероятность ошибки декодирования символа блочного кода, исправляющего t ошибок при использовании в канале без памяти, можно оценить следующим образом [26, 27]. Будем считать, что набор из $i > t$ ошибок в канале приводит к тому, что декодированное слово будет отличаться от переданного в $i+t$ позициях, так что $(i+t)/N_1$ -ая часть информационных символов декодируется неправильно. Полагая также в первом приближении, что при этом примерно половина битов, образующих символ кодового слова, будет декодирована неправильно, получаем, что вероятность ошибки на бит может быть оценена как:

$$P_b \leq \frac{1}{2} \sum_{i=t+1}^{N_1} \frac{i+t}{N_1} C_{N_1}^i p_e^i (1-p_e)^{N_1-i}, \quad (17)$$

где p_e — вероятность ошибки декодирования символа внутреннего кода по МП, граница для которой вычислена согласно (8) или (16).

Далее рассмотрим три конкретных примера каскадной конструкции (они представлены в таблице 2):

1) внутренний код построен на основе решетки E_8 , а внешний — это код Рида — Соломона с параметрами $N_1 = 64$, исправляющий $t = 1, 2, 3, 4$ ошибки с числом информационных символов $K_1 = 62, 60, 58, 56$ и с минимальным расстоянием $d_1 = N_1 - K_1 + 1 = 3, 5, 7, 9$;

2) внутренний код построен на основе решетки RE_8 , а внешний — это код Рида — Соломона с параметрами $N_1 = 256$, исправляющий $t = 1, 2, 3, 4$ ошибки, с числом информационных символов $K_1 = 254, 252, 250, 248$ и с минимальным расстоянием $d_1 = N_1 - K_1 + 1 = 3, 5, 7, 9$;

3) внутренний код построен на основе решетки Λ_{16} , а внешний — это код Рида — Соломона с параметрами $N_1 = 1024$, исправляющий $t = 1, 2, 3, 4$ ошибки, с числом информационных символов $K_1 = 1022, 1020, 1018, 1016$ и с минимальным расстоянием $d_1 = N_1 - K_1 + 1 = 3, 5, 7, 9$.

На рисунках 11-13 показаны зависимости вероятности ошибки на бит от числа циклов перезаписи для трех выбранных кодовых конструкций, полученные согласно выражению (17). Для всех трех примеров внешний код, исправляющий одну ошибку, *не обеспечивает* выполнения условия $P_b < 10^{-12}$ ни для одного из вариантов параметров канала (времени хранения T и числа циклов перезаписи N). Данные, представленные на рисунках 11-13, относятся к внешним кодам, которые могут обеспечить вероятность ошибки на бит $P_b < 10^{-12}$ хотя бы для некоторых значений хранения T и числа циклов перезаписи N .

Во всех рассмотренных случаях плотность записи оказывается умеренной. Из рассмотрения графиков можно сделать вывод о том, что при небольшом числе циклов перезаписи (примерно до 10^4) предпочтительным из рассмотренных вариантов оказывается конструкция $E_8 + (64, K_1)$. При этом обеспечивается плотность записи порядка 1.30...1.35 бит/ячейка. При увеличении числа циклов перезаписи примерно до $2 \cdot 10^4$ для сохранения уровня $P_b < 10^{-12}$ следует использовать конструкцию $\Lambda_{16} + (1024, K_1)$. При этом плотность записи снижается до величины порядка 1.24 бит/ячейка. При последующем уве-

личении числа циклов перезаписи конструкция $RE_8 + (256, K_1)$ может обеспечить выполнение условия $P_b < 10^{-12}$ при снижении плотности записи до уровня порядка 0.97 бит/ячейка.

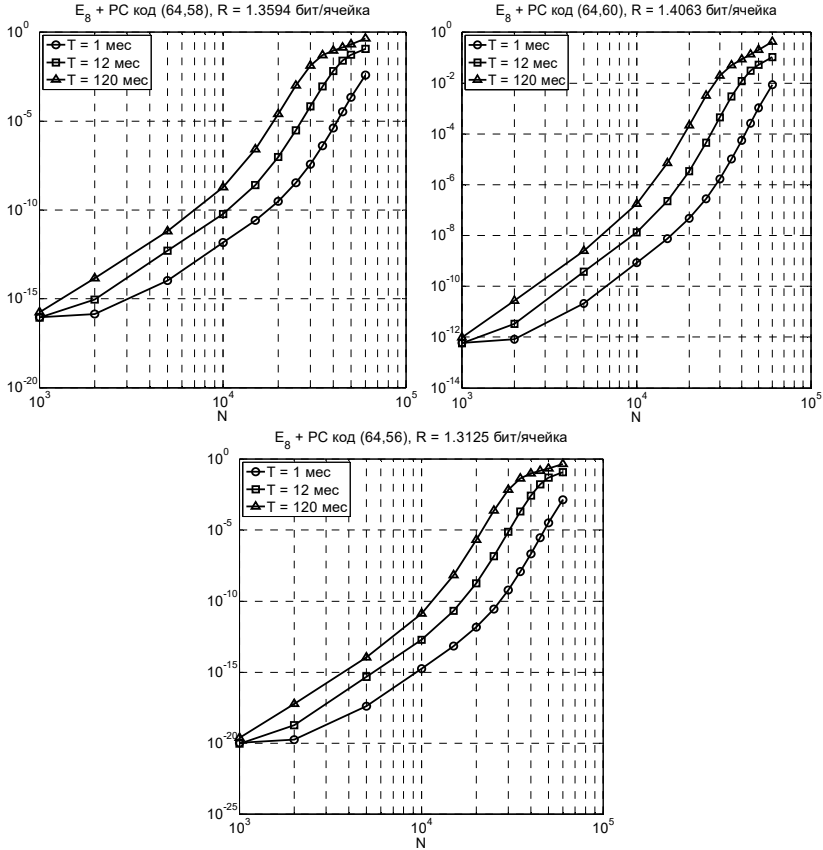


Рис. 11. Вероятность ошибки на бит для каскадных кодов $E_8 + (64, K_1)$, $K_1 = 60, 58, 56$, в зависимости от числа циклов перезаписи N

Если в качестве ключевого параметра, определяющего возможность использования рассматриваемой каскадной схемы для многоуровневой флэш-памяти, выступает сложность декодирования кода Рида — Соломона, то наименьшую сложность внешнего декодера (при $P_b < 10^{-12}$) среди проанализированных конструкций имеют

варианты (64,60), (256,252) и (1024,1020), позволяющие исправить 2 ошибки. При этом: а) плотность записи составляет 1.406, 0.984 и 1.245 бит/ячейка, соответственно; б) при времени хранения $T = 1$ мес. предельное число циклов перезаписи лежит в интервале от $2 \cdot 10^3$ до $2 \cdot 10^4$. Отметим, что согласно [28] сложность реализации декодера кода RS(256,252) составляет примерно 4 тысячи двухходовых вентиляей.

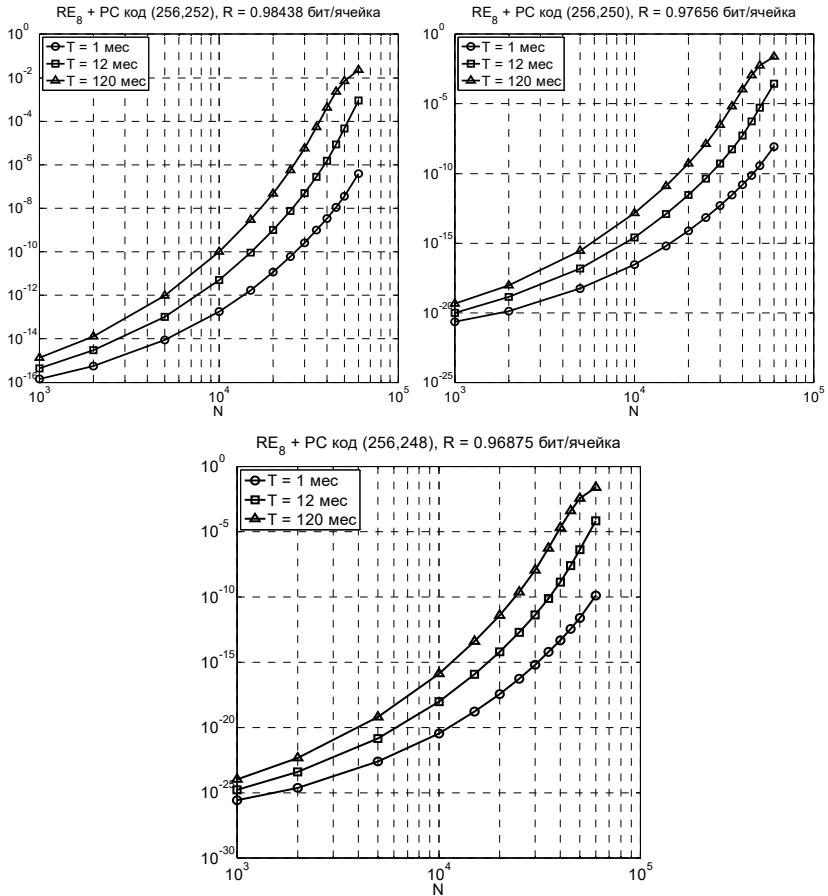


Рис. 12. Вероятность ошибки на бит для каскадных кодов $RE_b + (256, K_1)$, $K_1 = 252, 250, 248$, в зависимости от числа циклов перезаписи N

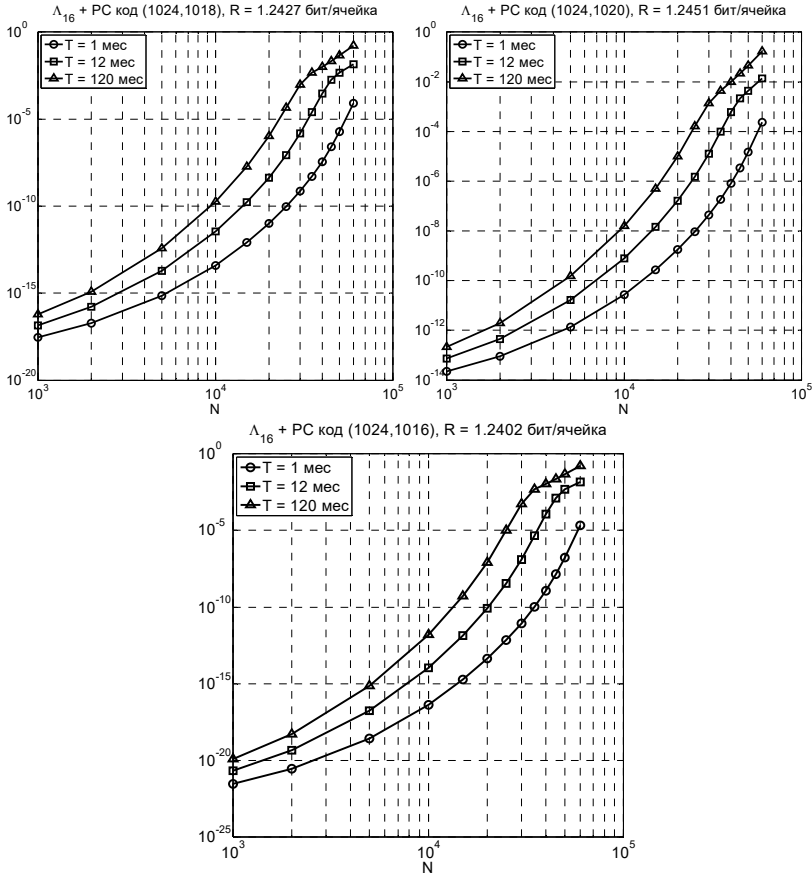


Рис. 13. Вероятность ошибки на бит для каскадных кодов $\Lambda_{16} + (1024, K_1)$, $K_1 = 1020, 1018, 1016$, в зависимости от числа циклов перезаписи

6. Заключение. Приближение плотности записи в многоуровневой флэш-памяти к теоретико-информационной границе [29], достигаемое с использованием помехоустойчивого кодирования, оборачивается, как правило, весьма высокой сложностью декодирования. Вместе с тем во многих практических приложениях в качестве ключевого параметра, определяющего возможность использования помехоустойчивого кодирования для многоуровневой флэш-памяти, выступает именно сложность декодирования. Для предложенной ранее весьма гибкой и достаточно эффективной каскадной конструкции на основе многомерной целочисленной решетки и кода Рида — Соломона [14], сложность

каскадного декодера практически полностью определяется сложностью декодирования внешнего кода Рида — Соломона, поэтому основной путь снижения сложности каскадного декодера — использование в рассматриваемой каскадной схеме кодов Рида — Соломона с небольшим числом исправляемых ошибок.

В настоящей работе была рассмотрена каскадная схема, в которой внешний код Рида — Соломона исправляет малое число ошибок — не более 4...5. В свою очередь, снижение корректирующей способности внешнего кода потребовало повышения помехоустойчивости внутренней ступени кодирования, реализуемого переходом к более широкому классу внутренних кодов, а именно кодов на основе решеток Барнса — Уолла E_8 , RE_8 , A_{16} и других. С учетом структуры решеток Барнса — Уолла, для оценки вероятности ошибки декодирования внутреннего кода была разработана модификация предложенного ранее в [14] оригинального подхода анализа помехоустойчивости кодирования в канале с аддитивным гауссовским шумом, дисперсия которого зависит от записанного значения (input-dependent additive Gaussian noise). Модификация основана на использовании параллельной структуры кодовой решетки внутреннего кода, что позволяет существенно понизить сложность вычислений и ускорить получение окончательного результата. Проведенный численный анализ показал следующее. Во-первых, в каскадных конструкциях с использованием рассмотренных решеток Барнса — Уолла для обеспечения вероятности ошибки $P_b < 10^{-12}$ код Рида — Соломона должен исправлять не менее двух ошибок. Во-вторых, по сравнению с ранее предложенными конструкциями [14], достижимая плотность записи умеренно снижается — на 22% ... 27%, в обмен на существенное (по меньшей мере, в 3...4 раза) уменьшение сложности каскадного декодера. Дальнейшее улучшение для рассмотренной конструкции обменного соотношения «плотность записи — сложность» возможно при использовании подходящего варианта мягкого декодирования кода Рида — Соломона.

Литература

1. *Michelsoni R., Marelli A., Ravasio R.* Error Correction Codes for Non-Volatile Memories // Springer Science & Business Media. 2008. 338 p.
2. *Choi H., Liu W., Sung W.* VLSI implementation of BCH error correction for multi-level cell NAND flash memory // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2010. vol. 18. no. 5. pp. 843–847.
3. *Freudenberger J., Spinner J.* A configurable Bose-Chaudhuri-Hocquenghem code architecture for flash controller applications // Journal of Circuits Systems and Computers. 2014. vol. 23. no. 02. pp. 1450019.
4. *Cho S., Kim D., Choi J., Ha J.* Block-wise concatenated BCH codes for NAND flash memories // IEEE Transactions on Communications. 2014. vol. 62. no. 4. pp. 1164–1177.

5. *Marelli A., Micheloni R.* BCH and LDPC Error Correction Codes for NAND Flash Memories // 3D Flash Memories. 2016. pp. 281–320.
6. *Sun F., Rose K., Zhang T.* On the Use of Strong BCH Codes for Improving Multilevel NAND Flash Memory Storage Capacity // IEEE Workshop on Signal Processing Systems (SiPS): Design and Implementation. 2006. 5 p.
7. *Haratsch E.F.* LDPC Code Concepts and Performance on High-Density Flash Memory // Proceedings of Flash Memory Summit. 2014. vol. 5. no. 5.2. pp. 5.5.
8. *Haymaker K., Kelley C.A.* Structured bit-interleaved LDPC codes for MLC flash memory // IEEE Journal on Selected Areas in Communications. 2014. vol. 32. no. 5. pp. 870–879.
9. *Lin X. et al.* Joint non-uniform detection and low-complexity decoding for multi-level cell NAND flash memory // 2016 IEEE International Conference on Consumer Electronics-China (ICCE-China). 2016. pp. 1–6.
10. *Oh J., Ha J., Moon J., Ungerboeck G.* RS-enhanced TCM for multilevel flash memories // IEEE Transactions on Communications. 2013. vol. 61. no. 5. pp. 1674–1683.
11. *Kurkoski B.M.* Coded modulation using lattices and Reed-Solomon codes, with applications to flash memories // IEEE Transactions on Selected Areas in Communications. 2014. vol. 32. no. 5. pp. 900–908.
12. *Spinner J., Rajab M., Freudenberger J.* Construction of high-rate generalized concatenated codes for applications in non-volatile flash memories // 2016 IEEE 8th International Memory Workshop (IMW). 2016. pp. 1–4.
13. *Spinner J., Freudenberger J., Shavgulidze S.* A Soft Input Decoding Algorithm for Generalized Concatenated Codes // IEEE Transactions on Communications. 2016. vol. 64. no. 9. pp. 3585–3595.
14. *Таубин Ф.А., Трофимов А.Н.* Каскадное кодирование на основе многомерных решеток и кодов Рида-Соломона для многоуровневой флэш-памяти // Труды СПИИРАН. 2018. Вып. 2(57). С. 75–103.
15. *Chatzigeorgiou I., Demosthenous A., Rodrigues M.R., Wassell I.J.* Performance-complexity tradeoff of convolutional codes for broadband fixed wireless access systems // IET Communications. 2010. vol. 4. no. 4. pp. 419–427.
16. *Lee H.* A high-speed low-complexity Reed-Solomon decoder for optical communications // IEEE Transactions on Circuits and Systems II: Express Briefs. 2005. vol. 52. no. 8. pp. 461–465.
17. *Yang C., Emre Y., Chakrabarti C.* Product code schemes for error correction in MLC NAND flash memories // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2012. vol. 20. no. 12. pp. 2302–2314.
18. *Samanta J. et al.* RS (255, 249). Codec Based on All Primitive Polynomials Over GF(2⁸) // Proceedings of International Conference on Communication, Devices, and Computing (ICCDC). 2017. pp. 69–81.
19. *Forney G.D.* Coset Codes. I. Introduction and Geometrical Classification // IEEE Transactions on Information Theory. 1988. vol. 34. no. 5. pp. 1152–1187.
20. *Taranalli V., Uchikawa H., Siegel P.H.* Channel models for multilevel cell flash memories based on empirical error analysis // IEEE Transactions on Communications. 2016. vol. 64. no. 8. pp. 3169–3181.
21. *Yassine H., Coon J., Ismail M., Fletcher H.* Towards an analytical model of NAND flash memory and the impact on channel decoding // 2016 IEEE International Conference on Communications (ICC). 2016. pp. 1–6.
22. *Korkotsides S., Bikas G., Eftaxiadis E., Antonakopoulos T.* BER analysis of MLC NAND Flash memories based on an asymmetric PAM model // 2014 6th International Symposium on Communications, Control and Signal Processing. 2014. pp. 558–561.
23. *Lin S., Kasami T., Fujiwara T., Fossorier M.* Trellises and Trellis-Based Decoding Algorithms for Linear Block Codes // Kluwer Academic Publishers. 1998. 288 p.

24. *Vardy A.* Trellis structure of codes // Handbook of Coding Theory. 1998. vol. 2. pp. 1985–2117.
25. *Calderbank A.R.* The Art of Signaling: Fifty Years of Coding Theory // IEEE Transactions on Information Theory. 1998. vol. 44. no 6. pp. 2561–2595.
26. *Кларк Дж., Кейн Дж.* Кодирование с исправлением ошибок в системах цифровой связи // М.: Радио и связь. 1987.
27. *Tomlinson M. et al.* Error-Correction Coding and Decoding // Springer. 2017. 522 p.
28. *Bentoutou Y.* Performance Comparison of Real Time EDAC Systems for Applications On-Board Small Satellites // International Journal of Computer, Electrical, Automation, Control and Information Engineering. 2011. vol. 5. no. 5. pp. 466–469.
29. *Трофимов А.Н., Таубин Ф.А.* Теоретико-информационный анализ многоуровневой flash-памяти. Часть 1. Модель канала и границы случайного кодирования. // Информационно-управляющие системы. 2016. Т. 81. № 2. С. 49–59.

Таубин Феликс Александрович — д-р техн. наук, профессор, кафедра аэрокосмических компьютерных и программных систем института аэрокосмических приборов и систем, Санкт-Петербургский государственный университет аэрокосмического приборостроения (СПбГУАП). Область научных интересов: цифровые системы связи, методы помехоустойчивого кодирования, широкополосные системы, беспроводная связь. Число научных публикаций — 94. ftaubin@yahoo.com; ул. Б. Морская, 67, 190000, Санкт-Петербург, Российская Федерация; р.т.: +7(812)494-70-51; факс: +7(812)494-70-51.

Трофимов Андрей Николаевич — канд. техн. наук, доцент, кафедра инфокоммуникационных систем и кафедра безопасности информационных систем института информационных систем и защиты информации, Санкт-Петербургский государственный университет аэрокосмического приборостроения (СПбГУАП). Область научных интересов: теория цифровой связи, теория информации, методы помехоустойчивого кодирования. Число научных публикаций — 52. andrei.trofimov@vu.spb.ru; ул. Б. Морская, 67, 190000, Санкт-Петербург, Российская Федерация; р.т.: +7(812)494-70-52; факс: +7(812)494-70-52.

Поддержка исследований. Работа выполнена при поддержке Министерства образования и науки Российской Федерации при выполнении научно-исследовательской работы в рамках проектной части государственного задания в сфере научной деятельности по заданию № 2.2716.2014/К и научно-исследовательской работы № 2.9214.2017/ БЧ в рамках базовой части государственного задания в сфере научной деятельности на 2017–2019 годы.

F.A. TAUBIN, A.N. TROFIMOV
**CONCATENATED CODING FOR MULTILEVEL FLASH
MEMORY WITH LOW ERROR CORRECTION CAPABILITIES IN
OUTER STAGE**

Taubin F.A., Trofimov A.N. Concatenated Coding for Multilevel Flash Memory with Low Error Correction Capabilities in Outer Stage.

Abstract. One of the approaches to organization of error correcting coding for multilevel flash memory is based on concatenated construction, in particular, on multidimensional lattices for inner coding. A characteristic feature of such structures is the dominance of the complexity of the outer decoder in the total decoder complexity. Therefore the concatenated construction with low-complexity outer decoder may be attractive since in practical applications the decoder complexity is the crucial limitation for the usage of the error correction coding.

We consider a concatenated coding scheme for multilevel flash memory with the Barnes-Wall lattice based codes as an inner code and the Reed-Solomon code with correction up to 4...5 errors as an outer one.

Performance analysis is fulfilled for a model characterizing the basic physical features of a flash memory cell with non-uniform target voltage levels and noise variance dependent on the recorded value (input-dependent additive Gaussian noise, ID-AGN). For this model we develop a modification of our approach for evaluation the error probability for the inner code. This modification uses the parallel structure of the inner code trellis which significantly reduces the computational complexity of the performance estimation. We present numerical examples of achievable recording density for the Reed-Solomon codes with correction up to four errors as the outer code for wide range of the retention time and number of write/read cycles.

Keywords: Concatenated Code, Barnes-Wall Lattices, Reed-Solomon Code, Multilevel Flash Memory, Decoding Error Probability, Decoding Complexity.

Taubin Feliks Aleksandrovich — Ph.D., Dr.Sci., Professor, Department of aerospace computer and software systems of Institute of Aerospace Instruments and Systems, Saint Petersburg State University of Aerospace Instrumentation (SUAI). Research interests: communication theory, error-correcting coding, spread spectrum systems, wireless communication. The number of publications — 94. ftaubin@yahoo.com; 67, B. Morskaya str., 190000, St.Petersburg, Russian Federation; office phone: +7(812)494-70-51; fax: +7(812)494-70-51.

Trofimov Andrey Nikolaevich — Ph.D., Associate Professor, Department of Information and Department of Communication Systems and Information Systems Security of Institute of Information Systems and Information Security, Saint Petersburg State University of Aerospace Instrumentation (SUAI). Research interests: communication theory, error-correcting coding, information theory. The number of publications — 52. andrei.trofimov@vu.spb.ru; 67, B. Morskaya str., 190000, St.Petersburg, Russian Federation; office phone: +7(812)494-70-52; fax: +7(812)494-70-52.

Acknowledgements. This research is supported by the Ministry of Education and Science of the Russian Federation in the framework of the project part of the government task in the field of scientific activity on assignment No. 2.2716.2014/K, and within the framework of the government task in the field of scientific activity for 2017–2019 on assignment No. 2.9214.2017/БН.

References

1. Micheloni R., Marelli A., Ravasio R. Error Correction Codes for Non-Volatile Memories. Springer Science & Business Media. 2008. 338 p.
2. Choi H., Liu W., Sung W. VLSI implementation of BCH error correction for multi-level cell NAND flash memory. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. 2010. vol. 18. no. 5. pp. 843–847.
3. Freudenberger J., Spinner J. A configurable Bose-Chaudhuri-Hocquenghem codec architecture for flash controller applications. *Journal of Circuits Systems and Computers*. 2014. vol. 23. no. 02. pp. 1450019.
4. Cho S., Kim D., Choi J., Ha J. Block-wise concatenated BCH codes for NAND flash memories. *IEEE Transactions on Communications*. 2014. vol. 62. no. 4. pp. 1164–1177.
5. Marelli A., Micheloni R. BCH and LDPC Error Correction Codes for NAND Flash Memories. 3D Flash Memories. 2016. pp. 281–320.
6. Sun F., Rose K., Zhang T. On the Use of Strong BCH Codes for Improving Multilevel NAND Flash Memory Storage Capacity. IEEE Workshop on Signal Processing Systems (SiPS): Design and Implementation. 2006. 5 p.
7. Haratsch E.F. LDPC Code Concepts and Performance on High-Density Flash Memory. Proceedings of Flash Memory Summit. 2014. vol. 5. no. 5.2. pp. 5.5.
8. Haymaker K., Kelley C.A. Structured bit-interleaved LDPC codes for MLC flash memory. *IEEE Journal on Selected Areas in Communications*. 2014. vol. 32. no. 5. pp. 870–879.
9. Lin X. et al. Joint non-uniform detection and low-complexity decoding for multi-level cell NAND flash memory. 2016 IEEE International Conference on Consumer Electronics-China (ICCE-China). 2016. pp. 1–6.
10. Oh J., Ha J., Moon J., Ungerboeck G. RS-enhanced TCM for multilevel flash memories. *IEEE Transactions on Communications*. 2013. vol. 61. no. 5. pp. 1674–1683.
11. Kurkoski B.M. Coded modulation using lattices and Reed-Solomon codes, with applications to flash memories. *IEEE Transactions on Selected Areas in Communications*. 2014. vol. 32. no. 5. pp. 900–908.
12. Spinner J., Rajab M., Freudenberger J. Construction of high-rate generalized concatenated codes for applications in non-volatile flash memories. 2016 IEEE 8th International Memory Workshop (IMW). 2016. pp. 1–4.
13. Spinner J., Freudenberger J., Shavgulidze S. A Soft Input Decoding Algorithm for Generalized Concatenated Codes. *IEEE Transactions on Communications*. 2016. vol. 64. no. 9. pp. 3585–3595.
14. Taubin F.A., Trofimov A.N. [Concatenated Reed-Solomon/lattice coding for multi-level flash memory]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2018. vol. 2(57). pp. 75–103. (In Russ.).
15. Chatzigeorgiou I., Demosthenous A., Rodrigues M.R., Wassell I.J. Performance-complexity tradeoff of convolutional codes for broadband fixed wireless access systems. *IET Communications*. 2010. vol. 4. no. 4. pp. 419–427.
16. Lee H. A high-speed low-complexity Reed-Solomon decoder for optical communications. *IEEE Transactions on Circuits and Systems II: Express Briefs*. 2005. vol. 52. no. 8. pp. 461–465.
17. Yang C., Emre Y., Chakrabarti C. Product code schemes for error correction in MLC NAND flash memories. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. 2012. vol. 20. no. 12. pp. 2302–2314.
18. Samanta J. et al. RS (255, 249). Codec Based on All Primitive Polynomials Over GF(2⁸). Proceedings of International Conference on Communication, Devices, and Computing (ICCDC). 2017. pp. 69–81.
19. Forney G.D. Coset Codes. I. Introduction and Geometrical Classification. *IEEE Transactions on Information Theory*. 1988. vol. 34. no. 5. pp. 1152–1187.

20. Taranalli V., Uchikawa H., Siegel P.H. Channel models for multilevel cell flash memories based on empirical error analysis. *IEEE Transactions on Communications*. 2016. vol. 64. no. 8. pp. 3169–3181.
21. Yassine H., Coon J., Ismail M., Fletcher H. Towards an analytical model of NAND flash memory and the impact on channel decoding. 2016 IEEE International Conference on Communications (ICC). 2016. pp. 1–6.
22. Korkotsides S., Bikas G., Eftaxiadis E., Antonakopoulos T. BER analysis of MLC NAND Flash memories based on an asymmetric PAM model. 2014 6th International Symposium on Communications, Control and Signal Processing. 2014. pp. 558–561.
23. Lin S., Kasami T., Fujiwara T., Fossorier M. Trellises and Trellis-Based Decoding Algorithms for Linear Block Codes. Kluwer Academic Publishers. 1998. 288 p.
24. Vardy A. Trellis structure of codes. *Handbook of Coding Theory*. 1998. vol. 2. pp. 1985–2117.
25. Calderbank A.R. The Art of Signaling: Fifty Years of Coding Theory. *IEEE Transactions on Information Theory*. 1998. vol. 44. no 6. pp. 2561–2595.
26. Clark Jr. G.C., Cain J.B. *Error-Correcting Coding for Digital Communications*. Plenum Press. 1982. 432 p. (Russ. ed.: Klark Dzh., Kejn Dzh. *Kodirovanie s ispravleniem oshibok v sistemah cifrovoj svyazi*. M.: Radio i svyaz'. 1987).
27. Tomlinson M. et al. Error-Correction Coding and Decoding. Springer. 2017. 522 p.
28. Bentoutou Y. Performance Comparison of Real Time EDAC Systems for Applications On-Board Small Satellites. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*. 2011. vol. 5. no. 5. pp. 466–469.
29. Trofimov A.N., Taubin F.A. [Information theory analysis of multilevel flash memory. Part 1: Channel model and random coding bounds]. *Informacionno-upravljajushhie sistemy – Information and control systems*. 2016. Issue 81. no. 2. pp. 49–59. (In Russ.).

А.В. ФЕДОРЧЕНКО, Е.В. ДОЙНИКОВА, И.В. КОТЕНКО
**АВТОМАТИЗИРОВАННОЕ ОПРЕДЕЛЕНИЕ АКТИВОВ И
ОЦЕНКА ИХ КРИТИЧНОСТИ ДЛЯ АНАЛИЗА
ЗАЩИЩЕННОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ**

Федорченко А.В., Дойникова Е.В., Котенко И.В. Автоматизированное определение активов и оценка их критичности для анализа защищенности информационных систем.

Аннотация. Цель исследования — разработка методики автоматизированного выделения активов информационной системы и сравнительной оценки уровня их критичности для последующей оценки защищенности анализируемой целевой инфраструктуры. Под активами в данном случае понимаются все информационно-технологические объекты целевой инфраструктуры. Размеры, разнородность, сложность взаимосвязей, распределенность и динамичность современных информационных систем затрудняют определение целевой инфраструктуры и критичности информационно-технологических активов для ее корректного функционирования. Автоматизированное и адаптивное определение состава информационно-технологических активов и связей между ними на основе выделения статичных и динамичных объектов изначально неопределенной инфраструктуры является достаточно сложной задачей. Ее предлагается решить за счет построения актуальной динамической модели отношений объектов целевой инфраструктуры с использованием разработанной методики, которая реализует подход на основе корреляции событий, происходящих в системе. Разработанная методика основана на статистическом анализе эмпирических данных о событиях в системе. Методика позволяет выделить основные типы объектов инфраструктуры, их характеристики и иерархию, основанную на частоте использования объектов, и, как следствие, отражающую их относительную критичность для функционирования системы. Для этого в работе вводятся показатели, характеризующие принадлежность свойств одному типу, совместное использование свойств, а также показатели динамичности, характеризующие вариативность свойств относительно друг друга. Результирующая модель используется для сравнительной оценки уровня критичности типов объектов системы. В работе описываются используемые входные данные и модели, а также методика определения типов и сравнения критичности активов системы. Приведены эксперименты, показывающие работоспособность методики на примере анализа журналов безопасности операционной системы Windows.

Ключевые слова: информационно-технологические активы, типы активов, критичность активов, статистический анализ данных, корреляция событий безопасности, ушерб, оценка защищенности.

1. Введение. Одним из важнейших аспектов при анализе защищенности информационной системы является выделение и оценка критичности используемых информационно-технологических активов. Практически все современные организации так или иначе используют информационные технологии для бизнеса, их объекты связаны различными каналами передачи данных и множеством семантических взаимоотношений, а системы жизнеобеспечения их активов могут быть сильно распределены. Под активами в данном случае понимаются все информационно-технологические объекты целевой инфраструк-

туры (в том числе файлы, пользователи, базы данных, приложения, сервисы, серверы, рабочие станции и др.). Инфраструктура может значительно меняться во времени, в ней могут появляться новые незарегистрированные хосты, программно-аппаратное обеспечение, уязвимости и ошибки конфигурации. В подобных условиях при достаточно масштабной архитектуре целевой информационной инфраструктуры экспертное определение активов и степени их критичности требует значительных временных и человеческих затрат.

Автоматизация выявления активов и иерархии связей между ними, а также оценка их критичности позволят более точно и детально строить модель рисков для анализа защищенности. Под критичностью в данном случае понимается важность объекта для функционирования информационной системы. Определение критичности всех информационно-технологических объектов также важно для задачи выбора контрмер, поскольку реализация мер защиты в масштабных системах со сложными взаимосвязями может привести к непредвиденному побочному ущербу для критичных активов за счет отключения или удаления вспомогательных активов. Однако данная задача является достаточно сложной. Сетевые сканеры позволяют определять различные объекты сети, такие как сервисы, используемые ими порты, хосты, на которых они развернуты, и узлы связи. Тем не менее подобные средства не дают возможности самостоятельно выделять различные типы статичных и динамичных объектов (например, таких объектов, как процессы, сессии, пользователи, привилегии, операционные системы и др.) и их иерархию, что препятствует получению актуальной динамической модели изначально неопределенной информационной системы. Вычисление и анализ показателей объектов инфраструктуры, с точки зрения авторов, является основой для автоматизированного выявления ее наиболее важных типов объектов, а также их конкретных экземпляров, то есть наиболее критичных активов информационной системы. Дальнейшее объединение динамической модели объектов инфраструктуры с условно-статическими типами данных безопасности (уязвимости, эксплойты, программно-аппаратное обеспечение, шаблоны атак, слабости и др.) позволит выявлять вредоносные, нелегитимные и аномальные объекты, проводить проактивный мониторинг состояния защищенности организации для предупреждающей выработки контрмер.

К настоящему моменту исследователями были разработаны различные методики автоматизированного определения распространения ущерба от атак в целевой инфраструктуре и критичности активов информационных систем. В частности, хорошо себя показали подходы, основанные на графах зависимостей сервисов. В том числе в [1] авто-

ры предлагают использовать зависимости между ресурсами системы для определения критичности неосновных активов и последующего выбора мер противодействия кибератакам [2, 3]. Однако они ограничиваются одним типом объектов (сервисами) и зачастую не рассматривают вопрос автоматизации выявления самих сервисов и зависимостей между ними, а также ориентированы на ручное определение наиболее важных активов и их связь с объектами инфраструктуры.

Таким образом, разработка методик автоматизированного определения активов информационной системы, иерархических связей между ними, то есть динамической модели инфраструктуры, и критичности этих активов является актуальной задачей для оценки защищенности в условиях неопределенности. В основе предлагаемой методики лежат методы корреляции событий, накапливаемых в различных журналах безопасности информационной системы. В [4] авторы описали общую идею применения методов корреляции для выявления типов объектов неопределенной инфраструктуры на основе анализа событий безопасности. В данной работе предлагается конкретная методика выявления типов объектов, а также определения их критичности. Различные виды анализа (статистический, структурный, динамический и др.) таких событий позволяют оценить критичность разных активов (объектов) и их типов с точки зрения функционирования целевой инфраструктуры и ее отдельных элементов с максимальной детализацией (от корневых до конечных объектов), которая ограничена лишь исходными данными. В этом случае подразумевается выделение иерархии типов активов и их наиболее значимых объектов.

Используемый подход корреляции событий позволяет выявить основные типы объектов инфраструктуры, их характеристики и иерархию путем анализа эмпирических данных, а также сравнивать критичность объектов на основе статических и динамических (статодинамических) показателей использования и времени жизни. Ввиду того, что подход основан на анализе эмпирических данных, он позволяет выявить только те объекты (активы) информационной системы, информация о которых наблюдалась в анализируемых событиях. Таким образом, точность работы подхода зависит от уровня детализации при журналировании событий.

Статья организована следующим образом. Во втором разделе рассмотрены релевантные исследования в области оценки критичности информационно-технологических активов и распространения ущерба от атак, автоматизированного определения конфигурации и корреляции событий безопасности. В третьем разделе описана разработанная методика выявления активов и оценки их критичности на основе предлагае-

мого подхода к корреляции событий. В четвертом разделе представлены проведенные эксперименты. В завершающем разделе приводится заключение, и описываются направления будущих исследований.

2. Релевантные работы. При анализе защищенности оценка критичности необходима для определения ущерба от кибератак, а также для определения побочного ущерба при реализации мер безопасности. В различных исследованиях критичность определяется по-разному, в том числе как важность актива для функционирования системы, как стоимость актива, как стоимость замены актива. Критичность может вычисляться как с использованием качественных, так и количественных показателей. Для определения ущерба, нанесенного критичным активам, довольно широко применяются графы зависимостей сервисов. Они позволяют отследить распространение ущерба от атак в информационной системе. Их целью является определение того, как та или иная уязвимость информационной инфраструктуры, заданной в виде сервис-ориентированной архитектуры, повлияет на деятельность организации. Граф зависимостей сервисов представляет собой множество сервисов (активов) компьютерной сети, связанных между собой в соответствии с тем, как свойства безопасности одного сервиса зависят от свойств безопасности другого. Учет распространения ущерба через зависимости сервисов позволяет отрегулировать затраты на безопасность, чтобы они не превысили возможный ущерб, обосновать их, а также не упустить важные уязвимости, которые могут привести к серьезным последствиям.

В [5] показатель ущерба от атаки рассчитывается на основе карты системы (граф зависимостей, объединяющий приоритетные ресурсы), иерархии ресурсов (группировка ресурсов по типам с выделением контрмер для каждого типа) и стоимостной модели (в которой ресурсам назначаются стоимости) как сумма стоимостей узлов системной карты, на которые атакующее действие повлияло негативно. В данном случае критичность ресурса определяется его стоимостью. В [6] для определения побочного ущерба при реагировании используется дерево зависимостей между ресурсами и предлагается показатель, отражающий стоимость снижения производительности сервиса в результате потери его доступности. То есть критичностью сервиса определяется его производительность. В [7] предлагается подход к определению уровня распространяемого ущерба с использованием графа зависимостей сервисов. В данном случае критичность актива (сервиса) определяется в зависимости от его стоимости на ранговой шкале. В [8] граф зависимостей сервисов используется для определения побочного ущерба при реагировании на атаки.

Указанные исследования позволяют определить, как распространяется ущерб от кибератак в системе, или определить критичность основных активов организации по тому, какой ущерб в итоге приносит нарушение их безопасности. Однако они нацелены именно на сервис-ориентированные архитектуры, и их целью не является автоматизированное определение сервисов системы (или других объектов информационной системы) и связей между ними. Исследование, представленное в настоящей статье, напротив, в большей степени направлено на выделение динамических объектов неопределенной инфраструктуры и последующее сравнение их критичности на основе показателя использования. То есть в данном исследовании под критичностью актива (информационно-технологического объекта) понимается сравнительная частота его использования в системе по сравнению с другими объектами (частота обращений других объектов к данному объекту).

Рассмотрим работы, связанные с автоматическим выделением объектов информационной системы. Основным средством автоматического определения информационной инфраструктуры сети являются сетевые сканеры. Средства сканирования сети делятся на активные, такие как Nmap [9] и Nessus [10], и пассивные, такие как Wireshark [11]. Они позволяют определить топологию сети и частично выявить сервисы сети путем сканирования портов, определения использующих их сервисов и обслуживающего программно-аппаратного обеспечения. Однако эти средства не обладают достаточной степенью точности и полноты информации, и, кроме того, не ставят своей задачей определение зависимостей между ресурсами сети. Впоследствии администраторы могут вручную дополнить собранную информацию. Кроме того, информация о сервисах системы может быть уже представлена в информации о конфигурации системы [12].

Сетевые сканеры позволяют определить объекты с заранее заданной структурой и не предназначены для определения всех типов динамических объектов, таких как пользовательские и системные процессы, сессии, их иерархии, или их роли. В данном исследовании предполагается разработать методику автоматизированного определения динамических объектов неопределенной инфраструктуры на основе анализа событий, что позволит иметь актуальную динамическую структуру информационной системы, а также автоматизированно сравнить критичность объектов информационной системы.

Развернутый обзор методик автоматизированного выявления зависимостей между сервисами дан в [13]. Автор выделяет запрос доступных знаний о зависимостях [14], например в системных файлах или в специально созданном хранилище конфигурации, использование

программного кода для определения зависимостей [15], пассивные методы идентификации [16] на основе анализа взаимодействия объектов, методы на основе нейросетей [17] и методы интеллектуального анализа данных [18, 19]. В [13] также отмечается, что задача все еще не решена полностью.

Описанные работы в основном учитывают объекты одного типа (сервисы). В то время как методика, разрабатываемая в данном исследовании, направлена на автоматизированное выделение различных типов объектов. Данное исследование ближе всего к работам, использующим интеллектуальный анализ данных. Однако в [18] для выделения объектов используются только частотные показатели типов событий, в то время как в данном исследовании применяются и другие характеристики исходной информации, такие как использование и вариативность свойств и их значений, что позволяет более точно выделить типы объектов. Исследование [19] очень близко к текущему, однако авторы исследуют цепочки событий, тогда как в настоящей статье вначале определяются иерархии объектов и их типы, а потом на следующем шаге осуществляется переход на анализ цепочек событий отдельных объектов. Это позволяет точнее сравнивать цепочки событий в будущем — делить события по принадлежности к объектам и строить связи между объектами. В [20] рассматривается корреляция событий и генерация правил (шаблонов) корреляции событий по эмпирическим данным, в то время как цель настоящего исследования состоит в выделении типов объектов сети. Методы, связанные с корреляцией событий, подробнее рассмотрены в следующем разделе для выделения наиболее подходящих с точки зрения поставленной в исследовании задачи.

Поскольку для решения двух описанных выше задач предлагается использовать аппарат корреляции, рассмотрим исследования, посвященные корреляции событий [21]. Изначально корреляция данных применялась в рамках систем обнаружения вторжений для выявления связей между сетевыми событиями с целью их агрегации и последующего обнаружения атак [22-24]. Однако корреляция событий, происходящих в информационной системе, кроме выявления инцидентов и предупреждений безопасности, может применяться для решения различных задач безопасности, в том числе для определения взаимосвязей между разнородной информацией безопасности, для группировки низкоуровневых событий в высокоуровневые мета-события, для выявления типов объектов информационной системы и связей между ними [25].

Наиболее популярным и простым в реализации методом корреляции является метод, основанный на правилах [26, 27]. Главный недостаток этого метода заключается в больших временных затратах на

задание правил администратором. Эффективность данного метода напрямую связана с квалификацией администратора. Многие методы, в том числе основанные на шаблонах (сценариях) [26, 28], графах [28-30], конечных автоматах [31], и другие, используют различные модели для отображения событий и связей между ними, но могут быть реализованы с использованием правил [32].

В настоящее время более перспективными представляются подходы к корреляции, основанные на самообучении [33, 34], такие как байесовские сети [26, 35], вероятностное исчисление событий [36-38], иммунные сети [35], искусственные нейронные сети [35, 39] и другие. Преимуществом данных подходов является возможность независимой (безусловной) корреляции событий с минимальными ручными настройками. Тем не менее для построения моделей обучения необходим предварительный анализ данных, который сложно автоматизировать. Кроме того, предъявляются требования к оценке адекватности и качества моделей, и исходная обучающая выборка должна быть достаточно полной.

Хотя рассмотренные методы применялись для обнаружения вторжений, а не для решения задачи выявления типов активов информационной системы и связей между ними, предполагается, что подобный анализ данных также применим и для решения поставленной в исследовании задачи. В настоящей статье показывается, что динамический анализ событий, происходящих в системе, и вычисление статодинамических показателей, в том числе частотных характеристик типов событий, вариативности значений свойств, парного использования и вариативности свойств, помогает выявить основные источники событий (объекты) и иерархические взаимосвязи между типами объектов. Разработанная методика корреляции на основе динамического анализа данных для автоматизированного выделения активов подробно описана в следующем разделе.

3. Методика выделения типов объектов и определение критичности активов. Разработанная методика предназначена для автоматизированного определения активов информационной системы, иерархических связей между ними, то есть динамической модели инфраструктуры и критичности этих активов на основе их использования в информационной системе. Методика объединяет несколько последовательных этапов:

1. Сбор и предварительная обработка входных данных. На данном этапе производится сбор данных о событиях из журналов информационной системы и их нормализация для корректного проведения дальнейшего анализа.

2. Выделение типов объектов. На этом этапе применяется статистический анализ собранных данных для выявления типов характери-

стик, типов объектов целевой инфраструктуры, иерархических связей между ними на основе статодинамических показателей, в том числе частотных характеристик типов событий, парного использования и вариативности свойств.

3. Определение критичности объектов и их выделенных типов на основе показателя использования свойств и их значений.

Обобщенная схема первых трех этапов методики, которым поставлены в соответствие уровни, приведена на рисунке 1.

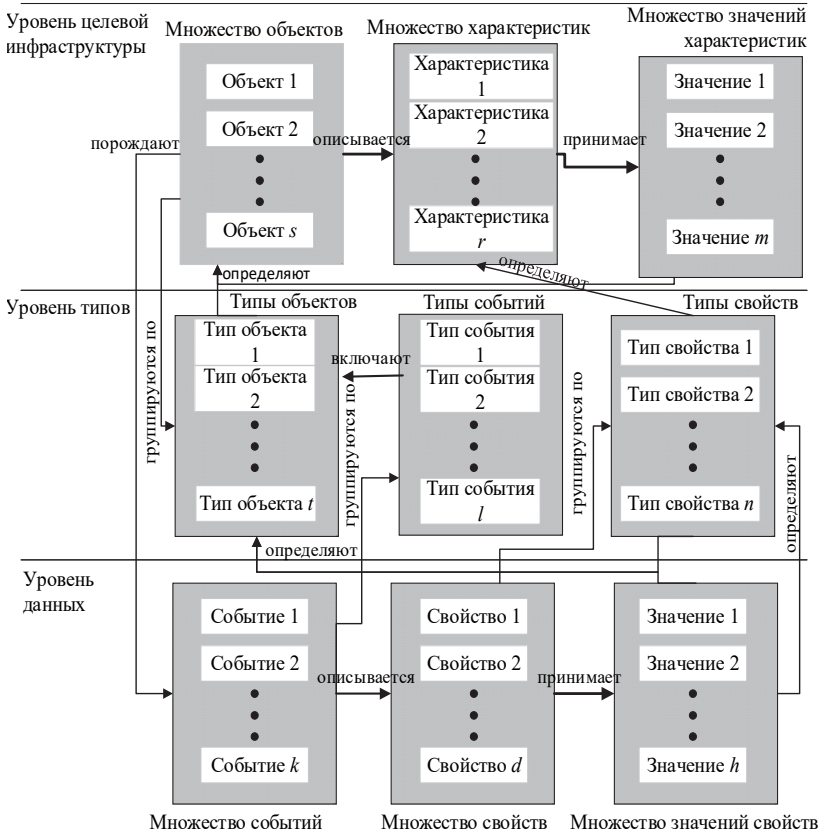


Рис. 1. Схема методики выделения типов объектов и их иерархии

Стрелками, сходящимися в одну точку, показана необходимость анализа нескольких сущностей для определения текущей (начало стрелки). Подробнее каждый этап описан в соответствующем подразделе.

3.1. Сбор и обработка входных данных. Предположим, что изначально анализируемая инфраструктура не определена. В качестве исходных данных используем множество событий, представленных в журналах объектов информационной системы. Рассмотрим подробнее исходные данные для выполнения процесса корреляции и основные понятия, используемые в рамках разработанной методики.

Исходными данными для выполнения статистического анализа является множество событий безопасности E журнала L :

$$E^L = \{e_1, e_2, \dots, e_n\}.$$

Под событием понимается факт либо результат выполнения какого-либо действия на любом из его этапов: попытка (событие отказа), начало действия (событие старта), промежуточный результат действия (события, продолжительные по времени), конечный результат выполнения действия (завершено корректно, завершено с ошибкой). Каждое событие e состоит из множества свойств p и их значений v , которыми описывается определенное действие [40]:

$$e = \{p_i, v_i\},$$

где p_i — свойство события, $p_i \in P^e$, $P^e \subset P$, P — множество всех свойств событий:

$$P = \{p_1, p_2, \dots, p_d\},$$

где d — общее количество свойств; v_i — значение свойства события, $v_i \in V^{p_i}$, V^{p_i} — множество значений свойства p_i :

$$V^{p_i} = \{v_1, v_2, \dots, v_h\},$$

h — количество возможных значений свойства p_i ; $V^{p_i} \subset V$, V — множество всех значений всех свойств:

$$V = \{V^{p_1}, V^{p_2}, \dots, V^{p_d}\}$$

В каждый момент времени данные множества являются конечными и не пустыми, хотя в течение времени они могут изменять свою мощность: $P \neq \emptyset$, $V^{p_i} \neq \emptyset$, $\forall i \in \{1, 2, \dots, d\}$.

Для корректного проведения дальнейшего анализа свойства событий и их значения должны быть нормализованы. В противном случае качество типизации свойств и выявление типов объектов могут быть существенно снижены. В рамках данного исследования не рассматривается задача автоматизации этого этапа, и задача выполняется на основе экспертной оценки.

Собранные данные делятся по типам принимаемых ими значений на категориальные и количественные. В зависимости от типа значений на следующем этапе используются разные подходы корреляции.

3.2. Выделение типов объектов и их иерархии. В работе для определения информационных объектов целевой инфраструктуры и их характеристик используются методы корреляции. Для описания подхода к корреляции событий на основе статистического анализа с целью выявления основных активов системы и иерархии связей между ними необходимо задать модель неопределенной инфраструктуры и определить ее связь с исходными данными.

При дискретно-непрерывном наблюдении (анализе истории наблюдений) целевая инфраструктура I , состав и архитектура которой изначально не определены, состоит из s информационных объектов O [40]:

$$O^I = \{o_1, o_2, \dots, o_s\},$$

имеющих некоторую продолжительность жизни. Их состояние описывается с помощью одной или нескольких характеристик x и их значений:

$$o = (x_i, v_i),$$

где $o \in O^I$, $x_i \in X$, $X = \{x_1, x_2, \dots, x_r\}$, $|X| \geq 1$, r — общее числа возможных характеристик объектов; $v_i \in V^{x_i}$, V^{x_i} — множество возможных значений характеристики x_i :

$$V^{x_i} = \{v_1, v_2, \dots, v_k\},$$

k — количество возможных значений характеристики x_i ; $V^{x_i} \subset V$, V — множество всех значений всех характеристик:

$$V = \{V^{x_1}, V^{x_2}, \dots, V^{x_k}\}.$$

Информационные объекты обязательно связаны друг с другом отношениями принадлежности. То есть каждый объект обязательно является частью более высокоуровневого объекта и (или) содержит в себе более низкоуровневые объекты. Также связь между объектами определяется за счет их непосредственного взаимодействия друг с другом. Это позволяет сформировать иерархию объектов.

Предполагается, что коллекция характеристик X^o , которыми описывается информационный объект o , однозначно задает тип информационного объекта ot , то есть каждая характеристика x из множества X содержится только в одном типе информационного объекта ot [40]. Типом объекта может быть «процесс» (сервис), «файл», «сессия» (сетевая, пользовательская), «сенсор», «хост» (сетевое оборудование, персональный компьютер), и другие.

Заключительным элементом модели неопределенной инфраструктуры I является множество отношений R между объектами O . Однако определение данного множества выходит за рамки настоящей работы и будет рассмотрено в дальнейших исследованиях.

Таким образом, модель M неопределенной инфраструктуры I состоит из множества информационных объектов O , а также множеств их типов OT и отношений OR :

$$M^I = \langle O, OT, OR \rangle.$$

Исходя из вышеописанного, методика выделения типов объектов и их иерархии должна преобразовывать исходные данные процесса корреляции к множествам объектов и их типов модели неопределенной инфраструктуры:

$$\{E, P, V\} \xrightarrow{f} \{O, OT\}.$$

Для выделения типа данных (смыслового) используется метод корреляции данных на основе динамического анализа событий. Этап выделения типов объектов и их иерархии включает следующие шаги: (1) выделение типов свойств событий на основе анализа их возможных значений; (2) выделение типов объектов на основе типов свойств; (3) определение иерархии типов объектов на основе показателя общего использования.

На первом шаге множество свойств $P = \{p_1, p_2, \dots, p_m\}$, где m — общее количество свойств всех событий, которые могут принимать

множество значений $V = \{V^{p_1}, V^{p_2}, \dots, V^{p_m}\}$, анализируется для выделения множества типов свойств pt , которое определяется как:

$$PT = \{pt_1, pt_2, \dots, pt_n\},$$

где n — общее количество типов.

Множества P и PT связаны отношением принадлежности, то есть каждый тип pt представляет подмножество свойств p :

$$pt = \{p_1, p_2, \dots, p_n\},$$

где l — количество свойств, описывающих тип pt .

Для определения принадлежности свойств определенному типу данных мы предлагаем использовать показатель парной вариативности их значений. Формула для вычисления данного показателя отличается для статичных и динамичных свойств. Под статичным понимается свойство, значение которого в среднем меняется намного реже, чем значение динамичного свойства (например, имя процесса), под динамичным понимается свойство, значение которого меняется несколько раз за то время, пока значение статичного свойства остается неизменным (например, идентификатор процесса). Для определения динамичности свойств на предмет изменения значений анализируются все значения всех свойств.

В роли показателей динамичности мы предлагаем использовать: (1) PV — абсолютное значение вариативности свойства, выраженное количеством возможных наблюдаемых (принимаемых) значений, и (2) EPV — средняя использование значений для каждого свойства, выраженная отношением количества событий, в которых наблюдается данное свойство, к общему числу возможных значений PV :

$$EPV(p) = |E^p| / PV. \quad (1)$$

В случае статичных свойств, показатель принадлежности свойств одному типу задается следующим образом:

$$\mu_{pt} = \frac{|V^{p_i} \cap V^{p_j}|}{|V^{p_i} \cup V^{p_j}|}, \quad (2)$$

где V — множество значений, принимаемых свойством, то есть принадлежность свойств p_i и p_j одному типу определяется эквивалентностью их значений.

В случае динамических свойств типы определяются путем анализа времени жизни и вариативности значений свойств на этом отрезке.

На втором шаге выполняется выделение групп свойств, предположительно характеризующих типы объектов на основе парного использования, и последующего уточнения этих групп за счет использования корреляции количественных значений и определения однотипных свойств внутри выделенных групп. Под парным использованием свойств p_i и p_j подразумевается отношение случаев их совместного использования к общему количеству их использований в событиях. Показатель парного использования между свойствами определяется по формуле:

$$\mu_{oi} = \frac{|E^{p_i} \cap E^{p_j}|}{|E^{p_i} \cup E^{p_j}|}, \quad (3)$$

Для дальнейшего уточнения выделенных типов объектов дополнительно используется определение однотипных свойств. Кроме того, возможно применение методов ранговой корреляции. Для этого значения свойства ранжируются по показателю использования в порядке убывания (отдельно для каждого свойства) и сопоставляются соответствующее его порядковому номеру.

На третьем шаге для определения типов объектов и их уровня в иерархии вычисляется значение общего использования их свойств (чем выше использование, тем выше объект в иерархии). Оно определяется как отношение количества использований свойства в событиях к общему количеству всех событий. Отношения между типами объектов определяются видами содержащих их событий. Можно выделить два основных вида: изменение состояния объекта и взаимодействие между объектами. Наличие взаимодействия определяет наличие связи между типами объектов.

3.3. Оценка критичности активов. Под критичностью актива (информационно-технологического объекта) будем понимать сравнительную частоту его использования в системе по сравнению с другими объектами.

Определение критичности объектов реализуется на основе выполнения следующих шагов:

- 1) Определение места типа объекта в иерархии типов объектов.
- 2) Определение относительной критичности объекта в соответствии с его местом в иерархии и по отношению к однотипным объектам.

Критичные объекты инфраструктуры предлагается определять как наиболее используемые объекты (показатель общего использования). На

первом шаге для определения места типа объекта в иерархии вычисляется показатель общего использования: $\mu = E/n$, где E — количество использования свойств типа объекта в событиях, n — общее количество событий. Типы объектов упорядочиваются сверху вниз, от максимально используемого к минимально используемому. Место l_{ot} , или уровень типа объекта ot в иерархии определяется как его порядковый номер.

На втором шаге критичность Cr определяется нормализацией уровня типа объекта в иерархии по общему количеству уровней l : $Cr = l_{ot}/l$. То есть имеется в виду относительная критичность информационно-технологического объекта определенного типа в системе.

Вышеописанные шаги позволяют определить уровень критичности относительно остальных типов (чем выше объект в иерархии, тем выше его использование, и, соответственно, тем выше его относительная критичность), то есть критичность типа объекта (актива) определяется относительным показателем его общего использования.

Необходимо отметить, что предлагаемый подход позволяет автоматизированно определить относительную критичность объектов инфраструктуры, однако требует дальнейшего усовершенствования и уточнения, в том числе за счет определения зависимостей между объектами и связанным с этим влиянием на критичность объектов. Кроме того, в данном случае не учитываются редко используемые критичные объекты, что также планируется рассмотреть в будущих исследованиях.

В дальнейшем планируется связать рассматриваемую динамическую инфраструктуру со статической, в том числе со стандартизированным определением программно-аппаратного обеспечения и уязвимостей. Это позволит связать предлагаемый показатель критичности объектов с вероятностью их компрометации и, следовательно, определять оценки рисков для объектов информационной системы.

4. Реализация и эксперименты. Реализация разработанной методики проводилась с использованием языка Python 3.5 и библиотек `pumpy`, `scipy` и `pandas`, а результаты были визуализированы с помощью `GraphViz` и модулей `matplotlib`, `pyplot` и `seaborn`.

Для выполнения экспериментов использовалась вычислительная платформа с одним 6-ядерным процессором Intel(R) Xeon(R) CPU E5-2603 v3 @ 1.60GHz. и 64 GB RAM.

4.1. Исходные данные для анализа. В условиях неопределенной целевой инфраструктуры множества событий E , свойств P и значений V изначально являются пустыми. Пользуясь историческими записями журналов событий либо накапливая информацию в реальном времени, следует первоначально определить множества свойств P и их возможных значений V .

Для проверки теоретических положений раздела 3 на данный момент эксперименты проводились в рамках одного хоста, в дальнейшем планируется провести тестирование для различных видов информационных систем. Исходными данными для проведения экспериментов являлись события системного журнала безопасности хоста под управлением ОС Windows 8. Подсистема журналирования операционной системы была настроена на сбор максимального количества типов событий безопасности. Анализируемый журнал обладает следующими характеристиками:

- количество событий ~ 6700000;
- количество оригинальных типов событий — 44 (из более чем 250 заявленных разработчиками для журнала безопасности [41]);
- количество свойств событий (включая свойства с единственным нулевым наблюдаемым значением) — 111;
- количество пар свойств, совместно наблюдаемых в событиях — 1379;
- количество возможных значений свойств без объединения повторяющихся вариантов между разными свойствами — 241024;
- общее количество возможных значений всех свойств — 213040;
- размер данных журнала — 7ГБ в формате XML, 1,25ГБ в формате CSV;
- время записи журнала — 36 дней;
- точность привязки событий к масштабу времени — десятые доли микросекунд.

На основе представленных характеристик анализируемого журнала можно сделать ряд предварительных выводов, а именно: (1) только 11% пар свойств (пункт 4) из возможного гипотетического количества пар наблюдаются в событиях совместно; (2) 12% возможных значений свойств наблюдаются как минимум в 2х свойствах (пункт 5 и 6); (3) количество событий в несколько десятков раз преобладает над количеством уникальных значений, что характеризует полноту журналирования поведения объектов инфраструктуры.

4.2. Предварительная обработка данных. Этап предварительной обработки исходной информации подразумевает нормализацию свойств событий, а также их значений для корректного проведения дальнейшего анализа (см. раздел 3). В настоящее время данная задача выполнена на основе экспертной оценки, а результаты нормализации описаны ниже.

Типы событий 5156 «The Windows Filtering Platform has allowed a connection» и 5157 «The Windows Filtering Platform has blocked a connection» содержат свойство «ProcessID», наименование которого

отличается от прочих свойств идентификатора процесса «ProcessId». Подобная малозначительная неточность может существенно снизить как качество типизации свойств, так и выявление типов объектов. С другой стороны, значительное отклонение структур свойств, описывающих информационные объекты, в данных типах событий может привести к неточным результатам анализа парного использования при однозначном сопоставлении указанных наименований. Возможным решением данной проблемы является приведение наименований свойств к однообразному регистру, а дальнейший анализ должен учитывать возможные пропуски в исходных данных при несогласованных структурах свойств типов событий. В описываемом случае наименование свойства «ProcessID» было приведено к «ProcessId».

В ходе анализа было установлено, что для 6 свойств событий наблюдалось единственное семантически-нулевое значение '-'. Такие свойства были исключены из дальнейшего анализа. Также указанное значение наблюдалось в 23 свойствах, имеющих прочие возможные значения, а само значение исключалось из списка возможных.

С нашей точки зрения, подобный пропуск в исходной информации может свидетельствовать о неверно составленной структуре типов событий, в которых данные свойства используются. Также для ряда свойств наблюдались единственные нулевые значения, представленные в разных форматах, например: «S-1-0-0» и «{00000000-0000-0000-0000-000000000000}». Однако данные значения несут определенную семантическую нагрузку, поэтому в дальнейшем анализе свойства, в которых они наблюдаются, были использованы.

4.3. Эксперименты. В рамках экспериментов, показатели, предложенные в разделе 3, использовались для выделения типов свойств (парная вариативность), типов объектов (парное использование) и их иерархии (общее использование свойств). Кроме того, были проведены эксперименты по определению динамичных свойств, позволяющих выделить критичные активы.

Первоочередным этапом анализа свойств событий является определение их показателей динамичности (предпосылки к выполнению данного этапа будут описаны ниже).

Как было определено в разделе 3, в роли таких показателей предлагается использовать PV (абсолютное значение вариативности свойства, выраженное количеством возможных наблюдаемых значений) и EPV (среднее использование значений для каждого свойства, определяемую по формуле 1).

Пример вычисленных показателей динамичности для некоторых наиболее показательных свойств представлен на рисунке 2.

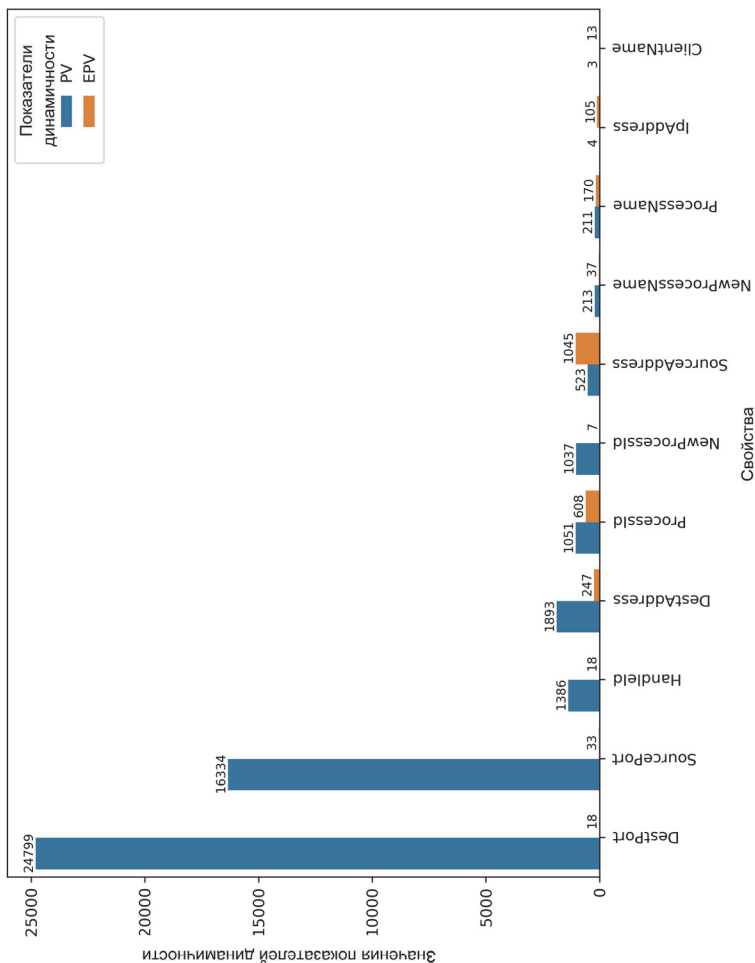


Рис. 2. Гистограмма показателей динамичности свойств событий

На данном рисунке видно, что наиболее динамичными являются свойства «DestPort» и «SourcePort». Однако такая оценка является достаточно грубой, поскольку динамичность свойства должна быть оценена во времени. Действительно, с уменьшением показателя абсолютной вариативности для некоторых свойств наблюдается рост среднего использования их значений, тогда как остальные свойства сохраняют тенденцию со значительным преобладанием вариативности.

Для подобных ситуаций следует гипотетически определить граничные значения показателей динамичности для:

(1) абсолютно статичного свойства, имеющего одно возможное значение;

(2) абсолютно динамичного свойства, все наблюдаемые значения которого являются уникальными в его пределах.

Указанные случаи описаны в таблице 1. Результаты измерений показателей динамичности некоторых свойств в течение времени (за 10 первых дней записей анализируемого журнала по дням) представлены на рисунке 3. На рисунке в каждый момент времени измеряются описанные показатели динамичности за весь прошедший период от начала наблюдений.

Таблица 1. Граничные значения

Вид свойства \ Показатель	PV	EPV
Статичное свойство	$PV \rightarrow 1$	$EPV \rightarrow E^p $
Динамичное свойство	$PV \rightarrow E^p $	$EPV \rightarrow 1$

Очевидно, что для большинства свойств показатель вариативности за рассматриваемый промежуток времени достигает определенных предельных значений. Данный факт не свидетельствует о достижении возможного максимума для последующих измерений, однако по данным графикам возможно оценить динамичность свойств по отношению друг к другу.

Явным недостатком предлагаемой оценки динамичности свойств являются абсолютные величины показателей. Другими словами, адекватно сравнивать динамичность свойств на основе данных показателей целесообразно при относительно высоком парном использовании свойств. В противном случае необходимо ввести относительную величину динамичности, что является одним из направлений дальнейших исследований.

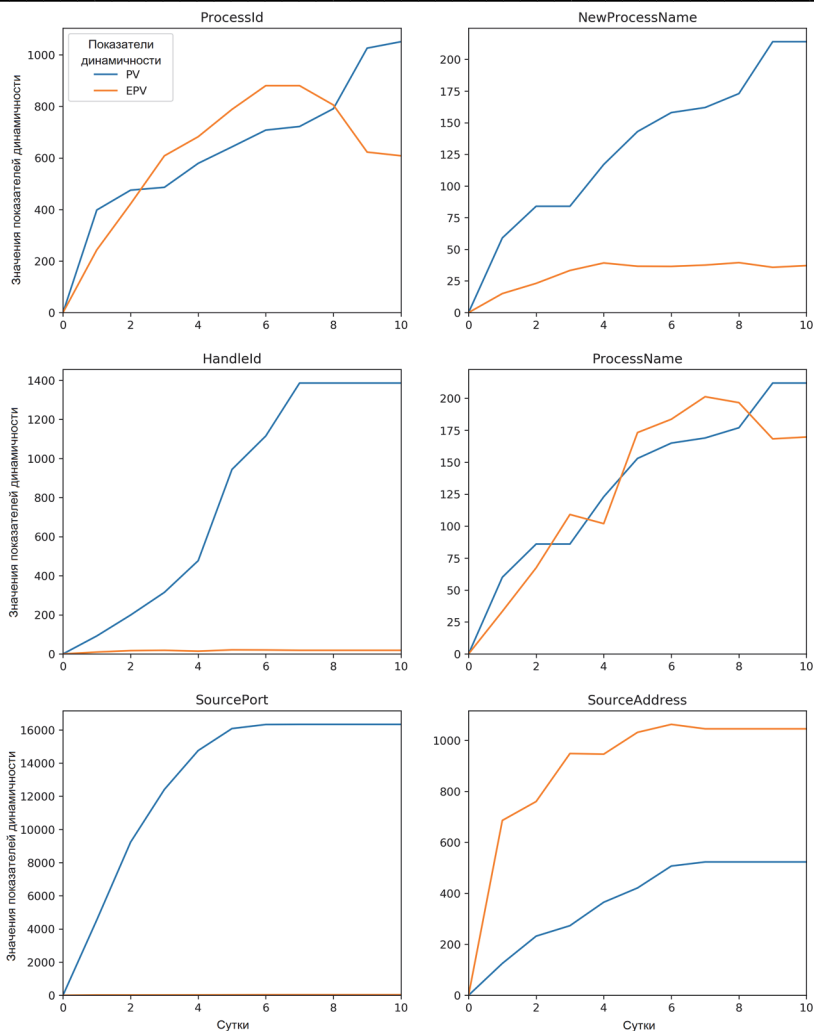


Рис. 3. Показатели динамичности свойств событий безопасности в течение времени за наблюдаемый период

4.4. Определение типов свойств. Задача выделения типов свойств событий, представленная в разделе 3, необходима для формирования списка характеристик информационных объектов, тогда как набор подобных характеристик будет в дальнейшем определять тип объекта описания. Для группировки свойств событий в типы используется показатель общей парной вариативности (формула 2).

Эксперименты показали, что не все значения показателей общей парной вариативности являются достаточными для автоматизированного принятия решения о принадлежности свойств к одному типу, что является ошибкой первого рода при исключении данной связи и, соответственно, нарушении верной типизации.

Причинами подобных ошибок могут являться: (1) высокая общая вариативность; (2) недостаточная общая вариативность одного из свойств. Однако более критичным случаем при типизации свойств является слияние двух разных типов событий, что является ошибкой второго рода при определении принадлежности разнотипных свойств к одному типу. В частности, такие ошибки могут быть вызваны схожестью форматов описания для разнотипных свойств.

Исходя из того, что вариативность является показателем динамичности свойства, преодоление описанных ошибок возможно при динамическом рассмотрении парной вариативности. Данная операция подразумевает сравнение возможных значений свойств в коротком интервале. Исходя из гипотезы, что при наблюдении определенного значения события в текущий момент времени, в окрестностях данного момента, объект, описываемый с помощью данного значения, является «живым». Таким образом, динамика изменения значений свойств событий определена временем жизни таких объектов. Парное сравнение значений свойств с привязкой к масштабу времени позволяет более точно определить парную вариативность свойств, и минимизировать появление описанных ошибок. В свою очередь, жизненный цикл объекта, описываемого с помощью конкретного значения свойства, может быть вычислен по диаграмме его наблюдения во времени.

Пример подобного графика представлен на рисунке 4.

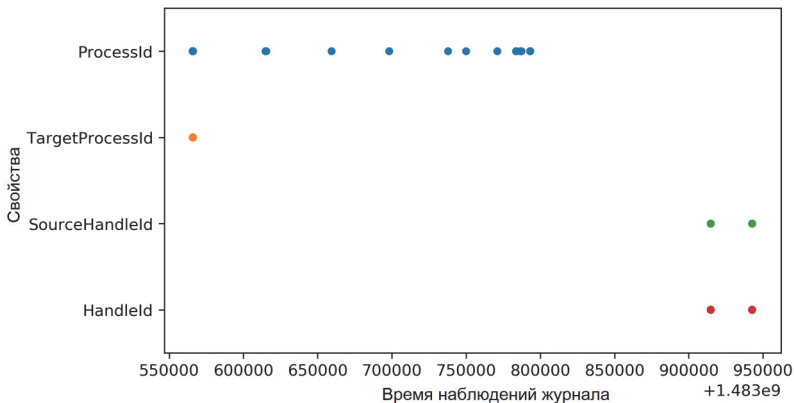


Рис. 4. Пример жизненных циклов объектов ОС Windows

По оси ординат на рисунке отображены свойства, которые принимают определенное значение (в данном случае «0x350»), а по оси абсцисс задано время наблюдения анализируемого журнала (36 дней) в относительных величинах.

Стоит отметить основное отличие между статичными и динамичными свойствами событий: статичные свойства характеризуют объект в общем масштабе времени, тогда как динамичное свойство характеризует объект лишь во время его жизни. Иными словами, при завершении жизненного цикла объекта динамическое свойство, которое его характеризовало, в последующих жизненных циклах будет принимать отличное от первоначального значение.

4.5. Определение типов объектов и их иерархии. Как было отмечено ранее, определение типов объектов по парному использованию свойств, которыми они описываются, основывается на гипотезе, что одинаковое использование свойств событий свидетельствует об описании характеристик одного или нескольких типов объектов одного уровня. Выдвигаемая гипотеза предварительно подтверждается проведенными экспериментами, показавшими, что отдельные группы свойств имеют равный либо очень близкий по значению показатель использования.

В ходе эксперимента по обнаружению разнотипных связей между свойствами на основе показателя их совместного использования в событиях было сформировано 18 групп свойств, а общее количество сгруппированных свойств равно 60. Стоит отметить, что свойства из одной группы используются в типах событий исключительно совместно [49].

Наиболее значимые и интерпретируемые типы объектов вместе со свойствами, которые их определяют, представлены на рисунке 5. На данном рисунке в каждой группе событий также указан показатель общего использования составляющих группу свойств, а сами группы упорядочены по данному показателю. Результаты проведенного эксперимента подтверждают гипотезу, согласно которой критичность типа объекта (актива) в ряде случаев определяется относительным показателем его общего использования (раздел 3).

Отдельно стоит рассмотреть объекты нулевого типа, которые встречаются во всех записях событий. Строго говоря, данные свойства определяют нулевой тип событий, так как даже такое пустое событие, как семантическая атомарная единица информации [42], имеет определенный заголовок со служебной информацией, например тип события «EventID» или задача «Task» [40].

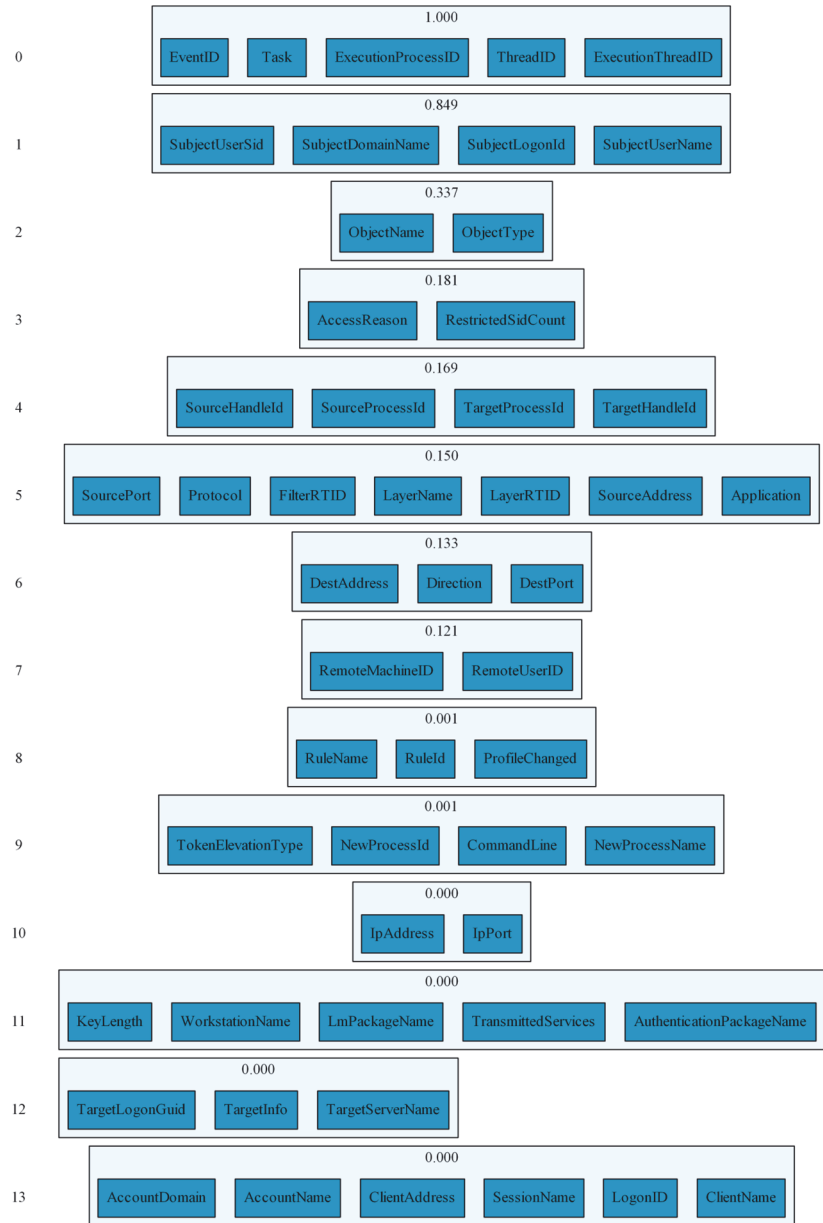


Рис. 5. Группы свойств ОС Windows, определяющие типы объектов, их уровень в иерархии и роли в типах событий

Рассмотрев сформированные группы объектов следует заключить, что большинство из них (за исключением нулевого уровня) являются достаточно семантически выраженными. Однако, представленные группы в действительности разбивают множество свойств не только по типам объектов, но и по их ролям в типах событий. Так, группы свойств 1 и 2 выделяют роли объектов «Субъект» и «Объект» соответственно.

В то же время группа 4 явно должна быть поделена на 2 роли «Источник» и «Цель» для одного типа объектов. Вероятно, более глубокая типизация объектов, например на основе вычисления коэффициентов корреляции, позволит избежать подобных неточностей. Данное направление будет рассматриваться в дальнейших исследованиях.

Отметим, что уровень в иерархии типов объектов, представленной на рисунке 5, определяет относительную критичность соответствующих активов по отношению к другим типам объектов (активов). В частности, группа характеристик 1, означающих соответствующие типы объектов (активов) «Домен» (*SubjectDomainName*), «Пользователь» (*SubjectUserName*, *SubjectUserSid*) и «Сессия» (*SubjectLogonId*) явно преобладает над остальными группам (кроме нулевой). Действительно, перечисленные типы объектов являются более критичными, чем типы: «Системный указатель» (*SourceHandleId*, *TargetHandleId*). «Процесс» (*SourceProcessId*, *TargetProcessId*), «Сетевой источник» (*DestPort*, *DestAddress*, *Direction*) и другие, что частично подтверждает выдвинутую гипотезу. Однако некоторые группы (7, 13), имеющие низкий показатель общего использования, явно не отражают уровень критичности соответствующих активов. Данный факт обусловлен редкостью наблюдения событий, связанных с объектами данных групп и корректная обработка подобных случаев является дальнейшим направлением исследований.

5. Заключение. В работе рассмотрен предлагаемый авторами подход к корреляции событий безопасности для оценки критичности активов (объектов) целевой инфраструктуры и их типов с точки зрения функционирования инфраструктуры и ее отдельных элементов с максимальной детализацией, которая ограничена лишь исходными данными. Описаны основные элементы подхода, в том числе исходные данные, модель неопределенной инфраструктуры, методика выделения типов объектов инфраструктуры и методика сравнения их критичности, а также применяемые на разных этапах показатели. Применение предложенного подхода показано на примерах и экспериментах, доказывающих выдвинутые теоретические предположения. Предлагается использовать разработанную методику оценки критичности в рамках оценки защищенности для определения возможного ущерба от реализации угроз безопасности. В дальнейшем планируется формализовать предлагаемый

подход, используя теорию нечетких множеств, усовершенствовать его за счет применения интервального анализа характеристик событий и расширить для выявления разных типов отношений между объектами, а также связать с методиками оценки защищенности. Кроме того, планируется развить методику оценки критичности, для учета случаев, когда критичность объекта не определяется тем, насколько часто его используют. Например, это касается генераторов ключей шифрования. Также планируется рассмотреть возможность улучшения методики оценки критичности за счет учета политики доступа организации.

Литература

1. *Kotenko I., Doynikova E., Chechulin A.* Security metrics based on attack graphs for the olympic games scenario // 2014 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing. 2014. pp. 561–568.
2. *Kotenko I., Doynikova E.* Countermeasure selection in SIEM systems based on the integrated complex of security metrics // 2015 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing. 2015. pp. 567–574.
3. *Doynikova E., Kotenko I.* Countermeasure selection based on the attack and service dependency graphs for security incident management // International Conference on Risks and Security of Internet and Systems. 2015. pp. 107–124.
4. *Kotenko I., Fedorchenko A., Saenko I., Kushnerevich A.* Parallelization of security event correlation based on accounting of event type links // 2018 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP). 2018. pp. 462–469.
5. *Balepin I., Maltsev S., Rowe J., Levitt K.* Using specification-based intrusion detection for automated response // International Workshop on Recent Advances in Intrusion Detection. 2003. pp. 136–154.
6. *Jahnke M., Thul C., Martini P.* Graph based metrics for intrusion response measures in computer networks // 32nd IEEE Conference on Local Computer Networks (LCN 2007) 2007. pp. 1035–1042.
7. *Kheir N. et al.* Cost Evaluation for Intrusion Response Using Dependency Graphs // 2009 International Conference on Network and Service Security. 2009. pp. 1–6.
8. *Shameli-Sendi A., Louafi H., He W., Cheriet M.* Dynamic Optimal Countermeasure Selection for Intrusion Response System // IEEE Transactions on Dependable and Secure Computing. 2018. vol. 15. no. 5. pp. 755–770.
9. NMap reference guide. URL: <http://nmap.org/book/man.html> (дата обращения: 02.07.2018).
10. Nessus vulnerability scanner. URL: <http://www.tenable.com/products/nessus-vulnerability-scanner> (дата обращения: 02.07.2018).
11. Wireshark vulnerability scanner. URL: <https://www.wireshark.org> (дата обращения: 02.07.2018).
12. *Clemm A., Bansal A.* Auto-Discovery at the Network and Service Management Layer // International Symposium on Integrated Network Management. 2003. pp. 365–378.
13. *Hanemann A.* Automated IT Service Fault Diagnosis Based on Event Correlation Techniques: Diss // Imu. 2007. 343 p.
14. *Steinder M., Sethi A.S.* A survey of fault localization techniques in computer networks // Science of Computer Programming. 2004. vol. 53. no. 2. pp. 165–194.
15. *Bagchi S., Kar G., Hellerstein J.* Dependency Analysis in Distributed Systems using Fault Injection: Application to Problem Determination in an e-commerce Environment // 12th International Workshop on Distributed Systems (DSOM'2001). 2001.

16. *Agarwal M.K. et al.* Mining Activity Data for Dynamic Dependency Discovery in e-Business Systems // IEEE Transactions on Network and Service Management. 2004. vol. 1. no. 2. pp. 49–58.
17. *Ensel C.* A scalable approach to automated service dependency modeling in heterogeneous environments // Proceedings Fifth IEEE International Enterprise Distributed Object Computing Conference. 2001. pp. 128–139.
18. *Tuchs K.D., Jobmann K.* Intelligent search for correlated alarm events in databases // 2001 IEEE/IFIP International Symposium on Integrated Network Management Proceedings. Integrated Network Management VII. Integrated Management Strategies for the New Millennium. 2001. pp. 285–288.
19. *Motahari-Nezhad H.R., Saint-Paul R., Casati F., Benatallah B.* Event correlation for process discovery from web service interaction logs // The VLDB Journal – The International Journal on Very Large Data Bases. vol. 20. no. 3. pp. 417–444.
20. *Hellerstein J.L., Ma S., Perng C.S.* Discovering actionable patterns in event data // IBM Systems Journal. 2002. vol. 41. no. 3. pp. 475–493.
21. *Федорченко А.В., Левиун Д.С., Чечулин А.А., Котенко И.В.* Анализ методов корреляции событий безопасности в SIEM-системах. Часть 1 // Труды СПИИРАН. 2016. Т. 4. № 47. С. 5–27.
22. *Artikis A. et al.* Scalable Proactive Event-Driven Decision Making // IEEE Technology and Society Magazine. 2014. vol. 33. no. 3. pp. 35–41.
23. *Raju B.K., Geethakumari G.* Event correlation in cloud: a forensic perspective // Computing. 2016. vol. 98. no. 11. pp. 1203–1224.
24. *Calyam P. et al.* Topology-Aware Correlated Network Anomaly Event Detection and Diagnosis // Journal of Network and Systems Management. 2014. vol. 22. № 2. pp. 208–234.
25. *Alevizos E. et al.* The Complex Event Recognition Group // ACM SIGMOD Record. 2018. vol. 47. no. 2. pp. 61–66.
26. *Sadoddin R., Ghorbani A.* Alert Correlation Survey: Framework and Techniques // Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services. 2006. pp. 37.
27. *Limmer T., Dressler F.* Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems // University of Erlangen, Dept. of Computer Science, Technical Report. 2008.
28. *Xu D., Ning P.* Correlation analysis of intrusion alerts // North Carolina State University. 2006.
29. *Michelioudakis E., Artikis A., Paliouras G.* Semi-Supervised Online Structure Learning for Composite Event Recognition // Machine Learning. 2018. pp. 1–26.
30. *Han Y., Zhu M., Liu C.* A Service-Oriented Approach to Modeling and Reusing Event Correlations // 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). 2018. vol. 1. pp. 498–507.
31. *Ghorbani A.A., Lu W., Tavallaee M.* Network Intrusion Detection and Prevention: Concepts and Techniques // Springer Science & Business Media. 2009. vol. 47. 223 p.
32. *Papataxiarhis V., Hadjiefthymiades S.* Event Correlation and Forecasting over Multivariate Streaming Sensor Data // arXiv preprint arXiv:1803.05636. 2018.
33. *Астахова Л.В., Цимбол В.И.* Применение самообучающейся системы корреляции событий информационной безопасности на основе нечеткой логики при автоматизации систем менеджмента информационной безопасности // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. 2016. Т. 16. № 1. 5 с.
34. *Tiwari R.R., Singh A.K., Singh V.* Self-learning SIEM system using association rule mining // Journal of Advanced Database Management & Systems. 2015. vol. 2. № 2. pp. 10–23.
35. *Gurer D.W., Khan I., Ogiev R., Keffer R.* An Artificial Intelligence Approach to Network Fault Management // SRI International. 1996. vol. 86.

36. *Skarlatidis A., Paliouras G., Artikis A., Vouros G.A.* Probabilistic Event Calculus for Event Recognition // ACM Transactions on Computational Logic (TOCL). 2015. vol. 16. no. 2. pp. 1–37.
37. *Alevizos E., Skarlatidis A., Artikis A., Paliouras G.* Probabilistic Complex Event Recognition: A Survey // ACM Computing Surveys. 2017. vol. 50. no. 5. pp. 71.
38. *Marvasti M.A., Poghosyan A.V., Harutyunyan A.N., Grigoryan N.M.* Statistical Normalcy Determination based on Data Categorization // VMware Technical Journal 2014. vol. 3. no. 1. pp. 43–55.
39. *Zhou J., Guo A., Celler B., Su S.* Fault detection and identification spanning multiple processes by integrating PCA with neural network // Applied Soft Computing. 2014. vol. 14. pp. 4–11.
40. *Федорченко А.В.* Анализ свойств событий безопасности для обнаружения информационных объектов и их типов в неопределенных инфраструктурах // Известия высших учебных заведений. Приборостроение. 2018. Вып. 61(11). С. 997–1004.
41. Windows Security Log Events. URL: <https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx> (дата обращения: 22.11.2018).
42. *Fedorchenko A., Kotenko I., El Baz D.* Correlation of security events based on the analysis of structures of event types // 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2017. vol. 1. pp. 270–276.

Федорченко Андрей Владимирович — младший научный сотрудник, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, обнаружение вторжений, вредоносные программы. Число научных публикаций — 40. fedorchenko@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7(812)328-7181.

Дойникова Елена Владимировна — канд. техн. наук, научный сотрудник, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, методы анализа рисков компьютерных сетей, управление информационными рисками. Число научных публикаций — 71. doynikova@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7(812)328-7181; факс: +7(812)328-4450.

Котенко Игорь Витальевич — д-р техн. наук, профессор, заведующий лабораторией, лаборатория проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, в том числе управление политиками безопасности, разграничение прав доступа, аутентификация, анализ защищенности, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей, анализ и верификация протоколов безопасности и систем защиты информации, защита программного обеспечения от взлома и управление цифровыми правами, технологии моделирования и визуализации для противодействия кибер-терроризму. Число научных публикаций — 450. ivkote@comsec.spb.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7(812)328-7181; факс: +7(812)328-4450.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проекты № 19-07-01246, 16-29-09482, 18-37-20047, 18-07-01488 и 18-29-22034), стипендии Президента РФ № СП-751.2018.5 и бюджетной темы АААА-А16-116033110102-5.

A. V. FEDORCHENKO, E. V. DOYNIKOVA, I. V. KOTENKO
**AUTOMATED DETECTION OF ASSETS AND CALCULATION OF
THEIR CRITICALITY FOR THE ANALYSIS OF INFORMATION
SYSTEM SECURITY**

Fedorchenko A.V., Doynikova E.V., Kotenko I.V. Automated Detection of Assets and Calculation of their Criticality for the Analysis of Information System Security.

Abstract. The research aims to develop the technique for an automated detection of information system assets and comparative assessment of their criticality for farther security analysis of the target infrastructure. The assets are all information and technology objects of the target infrastructure. The size, heterogeneity, complexity of interconnections, distribution and constant modification of the modern information systems complicate this task. An automated and adaptive determination of information and technology assets and connections between them based on the determination of the static and dynamic objects of the initially uncertain infrastructure is rather challenging problem. The paper proposes dynamic model of connections between objects of the target infrastructure and the technique for its building based on the event correlation approach. The developed technique is based on the statistical analysis of the empirical data on the system events. The technique allows determining main types of analysed infrastructure, their characteristics and hierarchy. The hierarchy is constructed considering the frequency of objects use, and as the result represents their relative criticality for the system operation. For the listed goals the indexes are introduced that determine belonging of properties to the same type, joint use of the properties, as well as dynamic indexes that characterize the variability of properties relative to each other. The resulting model is used for the initial comparative assessment of criticality for the system objects. The paper describes the input data, the developed models and proposed technique for the assets detection and comparison of their criticality. The experiments that demonstrate an application of the developed technique on the example of analyzing security logs of Windows operating system are provided.

Keywords: Assets, Asset Types, Asset Criticality, Statistical Data Analysis, Security Event Correlation, Impact, Security Assessment.

Fedorchenko Andrey Vladimirovich — junior researcher, Laboratory of Computer Security Problems, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: computer network security, intrusion detection, malware. The number of publications — 40. fedorchenko@comsec.spb.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russian Federation; office phone: +7(812)328-7181.

Doynikova Elena Vladimirovna — Ph.D., researcher, Laboratory of Computer Security Problems, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: computer network security, risk analysis methods for computer networks, information security risk management. The number of publications — 71. elenadoynikova@mail.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russian Federation; office phone: +7(812)328-7181; fax: +7(812)328-4450.

Kotenko Igor Vitalievich — Ph.D., Dr.Sci., Professor, Head of Laboratory, Laboratory of Computer Security Problems, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: computer network security, including security policy management, access control, authentication, network security analysis, intrusion detection, firewalls, deception systems, malware protection, verification of security systems, digital right management, modeling, simulation and visualization technologies for counteraction to cyber terrorism. The number of publications — 450. ivkote@comsec.spb.ru;

39, 14-th Line V.O., 199178, St. Petersburg, Russian Federation; office phone: +7(812)328-7181; fax: +7(812)328-4450.

Acknowledgements. This work was partially supported by grants of RFBR (projects No. 19-07-01246, 16-29-09482, 18-37-20047, 18-07-01488 and 18-29-22034), grant of the President of the Russian Federation SP-751.2018.5 and by the budget (the project No. AAAA-A16-116033110102-5).

References

1. Kotenko I., Doynikova E., Chechulin A. Security metrics based on attack graphs for th olympic games scenario. 2014 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing. 2014. pp. 561–568.
2. Kotenko I., Doynikova E. Countermeasure selection in SIEM systems based on the integrated complex of security metrics. 2015 23rd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing. 2015. pp. 567–574.
3. Doynikova E., Kotenko I. Countermeasure selection based on the attack and service dependency graphs for security incident management. International Conference on Risks and Security of Internet and Systems. 2015. pp. 107–124.
4. Kotenko I., Fedorchenko A., Saenko I., Kushnerevich A. Parallelization of security event correlation based on accounting of event type links. 2018 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP). 2018. pp. 462–469.
5. Balepin I., Maltsev S., Rowe J., Levitt K. Using specification-based intrusion detection for automated response. International Workshop on Recent Advances in Intrusion Detection. 2003. pp. 136–154.
6. Jahnke M., Thul C., Martini P. Graph based metrics for intrusion response measures in computer networks. 32nd IEEE Conference on Local Computer Networks (LCN 2007) 2007. pp. 1035–1042.
7. Kheir N. et al. Cost Evaluation for Intrusion Response Using Dependency Graphs. 2009 International Conference on Network and Service Security. 2009. pp. 1–6.
8. Shamel-Sendi A., Louafi H., He W., Cheriet M. Dynamic Optimal Countermeasure Selection for Intrusion Response System. *IEEE Transactions on Dependable and Secure Computing*. 2018. vol. 15. no. 5. pp. 755–770.
9. NMap reference guide. URL: <http://nmap.org/book/man.html> (дата обращения: 02.07.2018).
10. Nessus vulnerability scanner. Available at: <http://www.tenable.com/products/nessus-vulnerability-scanner> (accessed: 02.07.2018).
11. Wireshark vulnerability scanner. Available at: <https://www.wireshark.org> (accessed: 02.07.2018).
12. Clemm A., Bansal A. Auto-Discovery at the Network and Service Management Layer. International Symposium on Integrated Network Management. 2003. pp. 365–378.
13. Hanemann A. Automated IT Service Fault Diagnosis Based on Event Correlation Techniques: Diss. Imu. 2007. 343 p.
14. Steinder M., Sethi A.S. A survey of fault localization techniques in computer networks. *Science of Computer Programming*. 2004. vol. 53. no. 2. pp. 165–194.
15. Bagchi S., Kar G., Hellerstein J. Dependency Analysis in Distributed Systems using Fault Injection: Application to Problem Determination in an e-commerce Environment. 12th International Workshop on Distributed Systems (DSOM'2001). 2001.
16. Agarwal M.K. et al. Mining Activity Data for Dynamic Dependency Discovery in e-Business Systems. *IEEE Transactions on Network and Service Management*. 2004. vol. 1. no. 2. pp. 49–58.

17. Ensel C. A scalable approach to automated service dependency modeling in heterogeneous environments. Proceedings Fifth IEEE International Enterprise Distributed Object Computing Conference. 2001. pp. 128–139.
18. Tuchs K.D., Jobmann K. Intelligent search for correlated alarm events in databases. 2001 IEEE/IFIP International Symposium on Integrated Network Management Proceedings. Integrated Network Management VII. Integrated Management Strategies for the New Millennium. 2001. pp. 285–288.
19. Motahari-Nezhad H.R., Saint-Paul R., Casati F., Benatallah B. Event correlation for process discovery from web service interaction logs. *The VLDB Journal – The International Journal on Very Large Data Bases*. vol. 20. no. 3. pp. 417–444.
20. Hellerstein J.L., Ma S., Perng C.S. Discovering actionable patterns in event data. *IBM Systems Journal*. 2002. vol. 41. no. 3. pp. 475–493.
21. Fedorchenko A.V., Levshun D.S., Chechulin A.A., Kotenko I.V. [An Analysis of Security Event Correlation Techniques in Siem-Systems. Part 1]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2016. Issue 4. vol. 47. pp. 5–27 (In Russ.).
22. Artikis A. et al. Scalable Proactive Event-Driven Decision Making. *IEEE Technology and Society Magazine*. 2014. vol. 33. no. 3. pp. 35–41.
23. Raju B.K., Geethakumari G. Event correlation in cloud: a forensic perspective. *Computing*. 2016. vol. 98. no. 11. pp. 1203–1224.
24. Calyam P. et al. Topology-Aware Correlated Network Anomaly Event Detection and Diagnosis. *Journal of Network and Systems Management*. 2014. vol. 22. № 2. pp. 208–234.
25. Alevizos E. et al. The Complex Event Recognition Group. ACM SIGMOD Record. 2018. vol. 47. no. 2. pp. 61–66.
26. Sadoddin R., Ghorbani A. Alert Correlation Survey: Framework and Techniques. Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services. 2006. pp. 37.
27. Limmer T., Dressler F. Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems. University of Erlangen, Dept. of Computer Science, Technical Report. 2008.
28. Xu D., Ning P. Correlation analysis of intrusion alerts. North Carolina State University. 2006.
29. Michelioudakis E., Artikis A., Paliouras G. Semi-Supervised Online Structure Learning for Composite Event Recognition. *Machine Learning*. 2018. pp. 1–26.
30. Han Y., Zhu M., Liu C. A Service-Oriented Approach to Modeling and Reusing Event Correlations. 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). 2018. vol. 1. pp. 498–507.
31. Ghorbani A.A., Lu W., Tavallaee M. Network Intrusion Detection and Prevention: Concepts and Techniques. Springer Science & Business Media. 2009. vol. 47. 223 p.
32. Papataxiarhis V., Hadjiefthymiades S. Event Correlation and Forecasting over Multivariate Streaming Sensor Data. arXiv preprint arXiv:1803.05636. 2018.
33. Astahova L.V., Cimbol V.I. [Application of the self-learning system of information security events correlation based on the fuzzy logic for the information security management systems]. *Vestnik YUUrGU. Seriya: Kompyuternye tekhnologii, upravlenie, radioelektronika – Bulletin of the South Ural State University. Series: Computer Technologies, Automatic Control & Radioelectronics*. 2016. vol. 15 p. (In Russ.).
34. Tiwari R.R., Singh A.K., Singh V. Self-learning SIEM system using association rule mining. *Journal of Advanced Database Management & Systems*. 2015. vol. 2. № 2. pp. 10–23.
35. Gurer D.W., Khan I., Ogier R., Keffer R. An Artificial Intelligence Approach to Network Fault Management. SRI International. 1996. vol. 86.

36. Skarlatidis A., Paliouras G., Artikis A., Vouros G.A. Probabilistic Event Calculus for Event Recognition. *ACM Transactions on Computational Logic (TOCL)*. 2015. vol. 16. no. 2. pp. 1–37.
37. Alevizos E., Skarlatidis A., Artikis A., Paliouras G. Probabilistic Complex Event Recognition: A Survey. *ACM Computing Surveys*. 2017. vol. 50. no. 5. pp. 71.
38. Marvasti M.A., Poghosyan A.V., Harutyunyan A.N., Grigoryan N.M. Statistical Normalcy Determination based on Data Categorization. *VMware Technical Journal*. 2014. vol. 3. no. 1. pp. 43–55.
39. Zhou J., Guo A., Celler B., Su S. Fault detection and identification spanning multiple processes by integrating PCA with neural network. *Applied Soft Computing*. 2014. vol. 14. pp. 4–11.
40. Fedorchenko A.V. An [Analysis of Security Event Properties for Detection of the Information Objects and Their Types in Uncertain Infrastructures]. *Izvestiya vysshih uchebnyh zavedenij. Priborostroenie – Journal of Instrument Engineering*. 2018. vol. 61(11). pp. 997–1004. (In Russ.).
41. Windows Security Log Events. Available at: <https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx> (accessed: 22.11.2018).
42. Fedorchenko A., Kotenko I., El Baz D. Correlation of security events based on the analysis of structures of event types. 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). 2017. vol. 1. pp. 270–276.

И.В. ПРИСУХИНА, Д.В. БОРИСЕНКО, С.А. ЛУНЁВ
ИМИТАЦИОННАЯ МОДЕЛЬ ЭЛЕКТРИЧЕСКОГО КОДОВОГО СИГНАЛА В РОССИЙСКИХ СИСТЕМАХ ИНТЕРВАЛЬНОГО РЕГУЛИРОВАНИЯ ДВИЖЕНИЯ ПОЕЗДОВ НА ОСНОВЕ РЕЛЬСОВЫХ ЦЕПЕЙ

Присухина И.В., Борисенко Д.В., Лунёв С.А. Имитационная модель электрического кодового сигнала в российских системах интервального регулирования движения поездов на основе рельсовых цепей.

Аннотация. Системы интервального регулирования движения поездов на российских железных дорогах используют электрическую рельсовую цепь в качестве канала передачи информации о показании светофора на локомотив. Кодовые сигналы в таком канале подвержены влиянию помех, что часто приводит к нарушению движения поездов.

Существенного повышения помехоустойчивости можно достичь, если формализовать прием и дешифрацию сигналов электрической цепи в виде задачи классификации изображений. Для построения классификатора кодовых сигналов электрической рельсовой цепи с применением методов машинного обучения требуется обучающая выборка. Предлагается имитационная модель электрического кодового сигнала, позволяющая синтезировать такую выборку.

Имитационная модель структурирована в соответствии с основными этапами формирования и передачи кодового сигнала в рельсовой цепи: генератор кодового сигнала, рельсовая линия, приемное оборудование локомотива.

По результатам анализа осциллограмм и схем генераторов предложен алгоритм генерации кодового сигнала в начале рельсовой цепи. На этом этапе учтены временные характеристики кодовых сигналов, определяемые спецификацией, а также их случайные отклонения, обусловленные различными факторами.

Анализ схем замещения рельсовой линии, по которой передается кодовый сигнал, показал, что она представляет собой фильтр нижних частот. Для имитации влияния рельсовой линии на кодовый сигнал предложено использовать алгоритм цифрового фильтра Баттерворта. Параметры фильтра определены на основе электрических параметров рельсовой линии. Дополнительно на этом этапе учтено влияние случайных внешних помех.

Для имитации приемного оборудования локомотива, которое содержит полосовой фильтр, также предложено использовать алгоритм цифрового фильтра Баттерворта.

Таким образом, предложенная имитационная модель представляет собой совокупность последовательных алгоритмов. Изменяемые в заданных диапазонах параметры модели позволяют синтезировать осциллограммы кодовых сигналов с учетом различных условий работы компонентов электрической рельсовой цепи.

Ключевые слова: кодовый сигнал, автоматическая локомотивная сигнализация, машинное обучение, цифровой фильтр Баттерворта, рельсовая линия, нормальное распределение случайной величины, Octave.

1. Введение. Наиболее распространенными на сети железных дорог системами обеспечения безопасности движения поездов на переездах являются автоматическая блокировка числового кода (АБЧК) в совокупности с автоматической локомотивной сигнализацией непрерывного действия (АЛСН). Базовый принцип, лежащий в основе этих

систем, состоит в передаче информации о показании проходного светофора по рельсам посредством электрического сигнала, амплитудно-модулированного числовым кодом [1].

Накопленный опыт эксплуатации АБЧК и АЛСН показал, что серьезной проблемой является возникновение регулярных сбоев при дешифрации на локомотиве принятых из рельсов сигналов и, как следствие, нарушение бесперебойного движения поездов по перегону [2, 3]. Причины данных сбоев различны: влияние тягового тока и его гармоник; изменения в магнитном поле тягового тока; неравномерная намагниченность рельсов; пересечение рельсового пути линиями электропередач; искажения на этапах генерации сигналов (в приборах) и передачи по рельсовому тракту (в линии); электромагнитное влияние контактной подвески; разность высот локомотивных катушек над уровнем головки рельса и другие [2, 4-6].

Несмотря на широкое применение по всей сети железных дорог (порядка 67 тыс. километров), АБЧК и АЛСН морально устарели и уступают по ряду эксплуатационных показателей более совершенным своим модификациям (например, АБТЦ в сочетании с АЛС-ЕН и др. [2, 4]).

Однако, ввиду того что модернизация действующих участков сети дорог является целесообразной и экономически оправданной лишь при соответствующих объемах движения, скоростях и массах поездов, исследования, направленные на повышение устойчивости работы и качества функционирования АЛСН, являются по-прежнему актуальными [2].

Снизить число сбоев при дешифрации принятых на локомотиве сигналов в системе АЛСН, с сохранением принципа их формирования и рельсовой цепи как канала передачи можно за счет построения более совершенного дешифратора.

Как показано в [7-11], повышение качества дешифрации сигналов на локомотиве возможно за счет интеграции новых компонентов в состав его бортового оборудования, замены некоторых из них на более совершенные аналоги, использования методик цифровой обработки сигналов (например, корреляции сигналов).

Однако, с учетом возможностей и доступности современной вычислительной техники перспективным представляется подход, основанный на алгоритмах машинного обучения, в частности классификации изображений [12-16]. Преимуществом данного подхода на фоне, например, использования корреляции сигналов, является существенное уменьшение объема вычислительных операций при обработке очередного сигнала [17].

Так, например, больше нет необходимости в реализации процедур итеративного сравнения анализируемого сигнала с образцовыми, вычислении значений корреляционных функций, а также поиске максимального среди последних. В алгоритмах классификации изображений указанные мероприятия заменяются однократным расчетом используемой модели (например, искусственной нейронной сетью), что позволяет напрямую сформировать результат классификации на выходе системы без промежуточных вычислений.

Таким образом, предлагаемый подход обеспечивает увеличение скорости обработки сигналов на локомотиве, что, в свою очередь, способствует повышению безопасности движения поездов.

Алгоритм классификации сигналов, подлежащих распознаванию на локомотиве, строится на основе минимизации функции ошибки выбранной универсальной модели [18]. Минимизация функции ошибки подразумевает наличие некоторой совокупности сигналов, для которых заранее известен тип передаваемого ими числового кода. Данный процесс называется обучением, сигналы, для которых заранее известен тип числового кода, — обучающими примерами, а их совокупность — обучающей выборкой. Наличие обучающей выборки является необходимым условием построения алгоритма классификации.

Получение указанной выборки осциллограмм электрических сигналов, передающих числовой код, возможно двумя способами: с помощью регистрации осциллограмм на борту локомотива или их синтеза по имитационной модели электрических сигналов.

Императивным условием при обеспечении железнодорожных перевозок является безопасность следования поездов по участкам сети дорог. Строгое соблюдение данного требования в деятельности инфраструктурного и локомотивного комплексов ОАО «РЖД» исключает возможность проведения продолжительной регистрации таких осциллограмм, сопровождающейся подключением стороннего оборудования на борту локомотива. Таким образом, данный подход является бесперспективным для формирования обучающей выборки, так как лимитирует, во-первых, аудиторию ученых, имеющих доступ к зарегистрированным данным, а во-вторых, объем полученной выборки.

Альтернативным вариантом получения обучающей выборки может быть ее синтез по имитационной модели сигнала, передающего числовой код. Существенным преимуществом данного подхода по сравнению с предыдущим является его доступность для широкого академического сообщества, занятого вопросами машинной классификации электрических сигналов. Немаловажным достоинством выступает и легкостраи-

ваемый размер обучающей выборки, который регулируется при помощи выбора диапазонов изменения параметров имитационной модели.

Целью настоящей статьи выбрано рассмотрение вопросов, связанных с построением имитационной модели электрических сигналов, передающих числовой код.

Отметим, что вид обучающей выборки не зависит от выбранного алгоритма классификации (искусственная нейронная сеть, машина опорных векторов, LSTM-классификация и др. [13]), то есть он универсален. Это обстоятельство делает возможным разработку различных типов классификаторов на одной и той же обучающей выборке с последующим анализом их эффективности по критерию качества дешифрации сигналов на борту локомотива.

2. Структурная схема тракта передачи кодового сигнала. Электрические сигналы, передаваемые по рельсам на локомотив, в системе АЛСН для трехзначной системы АБЧК бывают трех видов: З (зеленый), Ж (желтый) и КЖ (красно-желтый). Данные сигналы предназначены для информирования машиниста о показании светофора, к которому приближается локомотив. При этом сигнал З соответствует зеленому огню на светофоре, Ж — желтому, а КЖ — красному. Эти сигналы на железнодорожном транспорте получили название «кодовые».

Тракт передачи кодовых сигналов З, Ж и КЖ на локомотив показан на рисунке 1 в виде укрупненных блоков (более детально тракт рассмотрен в [2]).

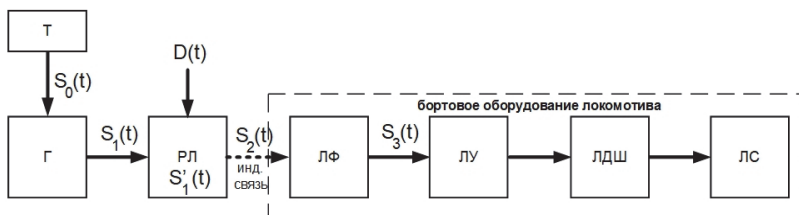


Рис. 1. Структурная схема тракта передачи кодовых сигналов З, Ж и КЖ на локомотив (Т – трансмиттер, Г – генератор, РЛ – рельсовая линия, ЛФ – локомотивный фильтр, ЛУ – локомотивный усилитель, ЛДШ – локомотивный дешифратор, ЛС – локомотивный светофор)

Кодовый сигнал $S_1(t)$ формируется генератором на основе сигнала числового кода $S_0(t)$, который поступает из трансмиттера (Т). При распространении по рельсовой линии (РЛ) с учетом ее влияния сигнал $S_1(t)$ преобразуется в сигнал $S'_1(t)$. В рельсовой линии на сигнал $S'_1(t)$ воздействует сигнал помехи $D(t)$, и на локомотив в месте нахождения его первой колесной пары посредством индуктивного взаимодействия

передается сигнал $S_2(t)$. Бортовое оборудование локомотива содержит локомотивный фильтр (ЛФ), локомотивный усилитель (ЛУ), локомотивный дешифратор (ЛДШ) и локомотивный светофор (ЛС). Сигнал $S_2(t)$ за счет влияния локомотивного фильтра преобразуется в $S_3(t)$.

Осциллограмма реального сигнала 3 $S_3^P(t)$, зафиксированного на выходе локомотивного фильтра на борту локомотива, показана на рисунке 2.

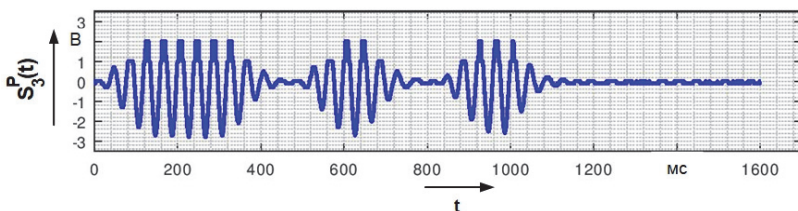


Рис. 2. Осциллограмма реального сигнала 3 на выходе локомотивного фильтра

По внешнему виду сигнала на рисунке 2 можно заключить, что он является амплитудно-модулированным синусоидальным. Также отчетливо видно, что сигнал имеет искажения формы.

В структуре имитационной модели кодовых сигналов предлагается выделить и отдельно рассмотреть три части:

- модулирующий сигнал;
- несущий сигнал;
- искажения формы кодового сигнала.

Средой моделирования авторами выбрана вычислительная платформа Octave 4.4.1.

3. Моделирование модулирующего сигнала. В качестве правила для амплитудной модуляции передаваемых по рельсам сигналов используется числовой код [5]. Сигналы числового кода (модулирующие сигналы) представляют собой циклические последовательности прямоугольных импульсов и интервалов. Число импульсов и их длительность определяются содержанием передаваемого информационного сообщения о показании светофора. Таким образом, для каждого показания светофора предусмотрена конкретная последовательность прямоугольных импульсов и интервалов [4].

Время цикла сигнала числового кода определяется типом кодового путевого трансмиттера (КПТШ), формирующего последовательность импульсов и интервалов заданной длительности [5]. В качестве примера выбраны КПТШ типов «5», «515», «8», «815», цикл кода которых составляет 1,6 с. Спецификация (номинальные длительности) генерируемых ими сигналов числового кода приведена в таблице 1.

Таблица 1. Номинальные длительности сигналов числового года, генерируемых КПТШ-5, КПТШ-515, КПТШ-8, КПТШ-815 [18]

Вид кодового сигнала	Длительность, мс					
	1 импульс	1 интервал	2 импульс	2 интервал	3 импульс	3 интервал
З	350	120	220	120	220	570
Ж	380	120	380	720	–	–
КЖ	230	570	230	570	–	–

Графики сигналов числового кода $S_0^3(t)$, $S_0^Ж(t)$, $S_0^{КЖ}(t)$, полученные в результате моделирования, согласно их спецификации (таблица 1) в среде Octave, представлены на рисунке 3. Частота дискретизации F_d выбрана равной 1000 Гц, а амплитуда — 1 В.

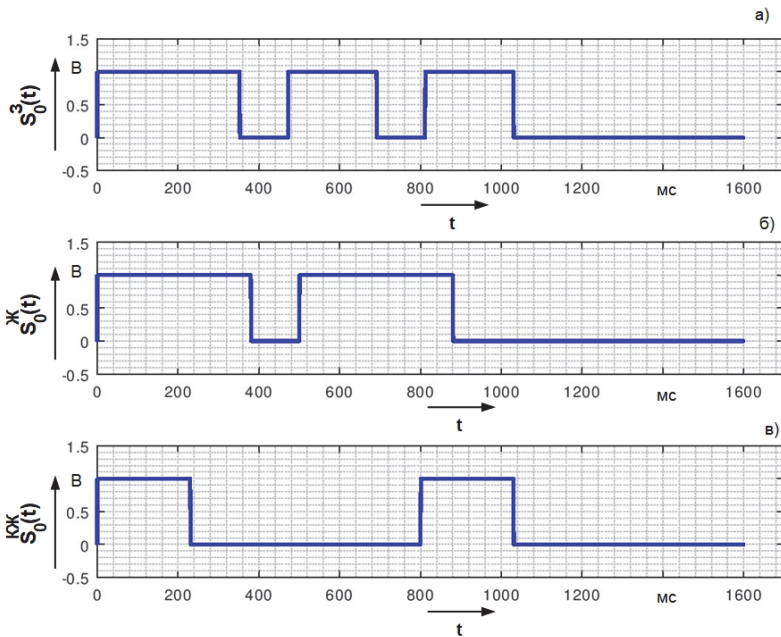


Рис. 3. Графики сигналов числового кода (а – код З, б – код Ж, в – код КЖ)

В реальных условиях временные параметры импульсов и интервалов сигналов числового кода имеют отклонения от номинальных значений. Это обусловлено воздействием ряда факторов [2, 4]. К наиболее существенным из них можно отнести следующие: разброс параметров и износ вращающихся деталей КПТШ, колебания напряжения его питания, инерционность работы и разброс параметров

транзиттерного реле, влияние аппаратуры рельсовой цепи, переходные процессы в локомотивном приёмнике.

Колебания напряжения питания КППШ оказывают негативное влияние на работу расположенного внутри асинхронного электродвигателя, что может стать причиной отклонения частоты его вращения от номинальной. Помимо этого, в КППШ с течением временем происходит постепенная деформация расположенных на валу кулачковых шайб асинхронного электродвигателя, которая протекает в зоне их взаимодействия с подвижным контактом.

Транзиттерное реле, отвечающее за коммутацию кодового сигнала в рельсовую цепь, в зависимости от условий эксплуатации имеет различную скорость переброса фронтового и тылового контактов, что приводит к инерционности его работы.

Поскольку на сети железных дорог применяются рельсовые цепи различных видов, состав аппаратуры их питающих и релейных концов не является однотипным. Как следствие, если аппаратура рельсовой цепи содержит устройства, которые имеют резонансную частоту, равную частоте несущего сигнала (например, защитный блок-фильтр, цепи дроссель-трансформаторов и др.), то нарастание тока в рельсах происходит медленнее, чем это предусмотрено номинальными характеристиками кодовых сигналов.

Нестабильность сопротивления балласта рельсовой цепи вследствие зависимости от погодных условий изменяет скорость распространения кодового сигнала по рельсам.

Негативное влияние при приеме кодового сигнала на борту локомотива оказывают значительные изменения его амплитуды под приемными катушками в пределах одной рельсовой цепи. При этом отношение крайних амплитуд (максимальной и минимальной) кодового сигнала может достигать значений 10 – 15. При несвоевременной регулировке чувствительности бортового приемника локомотива, указанные изменения амплитуды приводят к его некорректной работе и сбоям при дешифрации кодового сигнала.

Также, очевидно, что изменениям длительности импульсов соответствуют противоположные по длительности изменения интервалов.

Для имитации отклонений параметров сигналов числового кода от номинальных предлагается рассматривать положение границ импульсов и интервалов внутри их циклов как случайные величины, имеющие нормальное распределение. Так для сигнала З таких границ в цикле будет шесть, для Ж и КЖ — четыре.

Случайное положение границ импульсов и интервалов внутри цикла сигнала задается функциями f_i согласно (1) для кода З и (2) — для Ж и КЖ:

$$\{f_i \sim N(\mu_i, \sigma^2) \mid i \in [1, 6] \wedge i \in \mathbb{N}\}; \quad (1)$$

$$\{f_i \sim N(\mu_i, \sigma^2) \mid i \in [1, 4] \wedge i \in \mathbb{N}\}, \quad (2)$$

где i — номер границы внутри сигнала; $N(\mu_i, \sigma^2)$ — обозначение нормального распределения случайной величины; μ_i — математическое ожидание для i -й границы; σ^2 — дисперсия.

Для каждой функции f_i в качестве μ_i используются соответствующие номинальные значения из таблицы 2, отражающие положение границ импульсов и интервалов внутри цикла сигнала числового кода. Дисперсия σ^2 является изменяемым параметром имитационной модели.

Таблица 2. Номинальное положение границ импульсов и интервалов внутри циклов сигналов числового кода

μ_i	Положение границы внутри сигнала, мс		
	З	Ж	КЖ
1	350	380	230
2	470	500	800
3	690	880	1030
4	810	1600	1600
5	1030	—	—
6	1600	—	—

Данные рассуждения могут быть проиллюстрированы рисунком 4 на примере сигнала числового кода 3 ($S_0^3(t)$) единичной амплитуды. Функции $f_1 - f_6$ получены согласно (1).

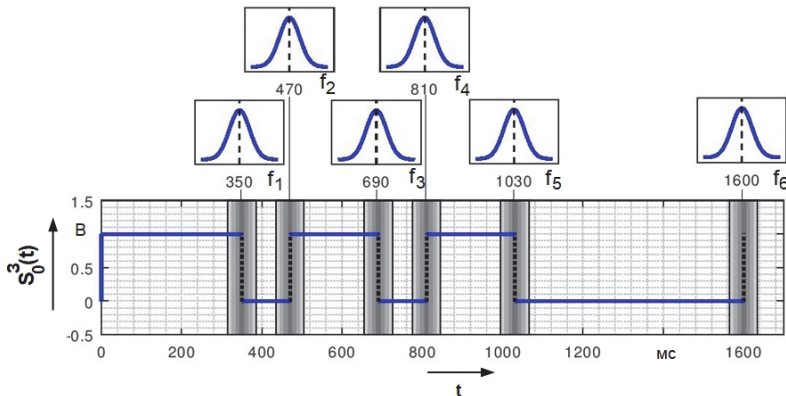


Рис. 4. Случайное положение границ импульсов и интервалов внутри сигнала числового кода 3

Графическая схема алгоритма (ГСА) генерации сигнала числового кода 3 показана на рисунке 5. Алгоритмы генерации сигналов числовых кодов Ж и КЖ описываются аналогичными блок-схемами.

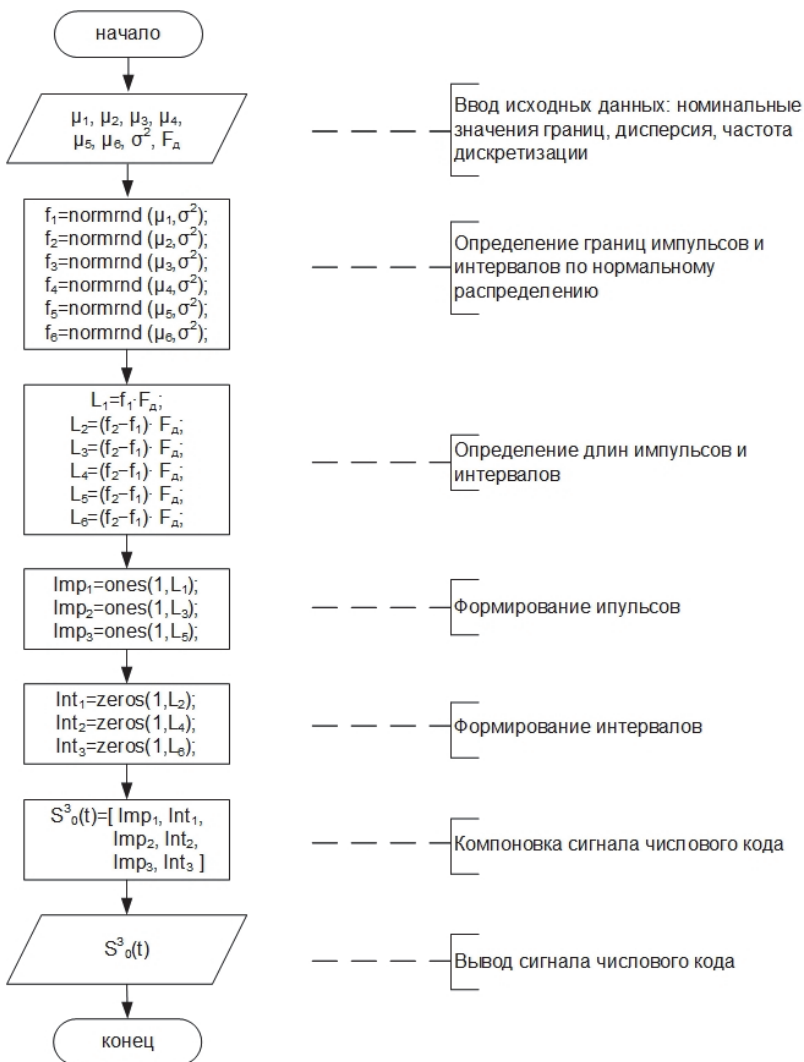


Рис. 5. ГСА генерации сигнала числового кода 3

4. Моделирование несущего сигнала. В качестве несущих в системе АЛСН используются синусоидальные сигналы частотой 25

или 50 Гц. Эти сигналы формируются специальными трансформаторами или статическими преобразователями частоты [19]. Фактические параметры несущих сигналов на выходе этих устройств зависят от ряда факторов, характерных для любой электротехнической системы: колебания питающего напряжения, его несинусоидальность, наличие в спектре напряжения высших гармонических составляющих, девиация частоты и другие. Однако данные отклонения в большинстве случаев являются несущественными на фоне помех, возникающих при передаче кодовых сигналов на локомотив. По этой причине данные отклонения предлагается отнести к допущениям имитационной модели, а несущий синусоидальный сигнал принять идеальным.

На рисунке 6 в качестве примера изображен амплитудно-модулированный числовым кодом 3 синусоидальный сигнал $S_1^3(t)$ единичной амплитуды с частотой несущего сигнала 25 Гц, полученный по формуле (3). Для унификации формул (здесь и далее) верхние индексы их принадлежности к конкретному числовому коду опущены.

$$S_1(t) = S_0(t) \cdot \sin(2\pi F_N t), \quad (3)$$

где F_N — частота несущего сигнала.

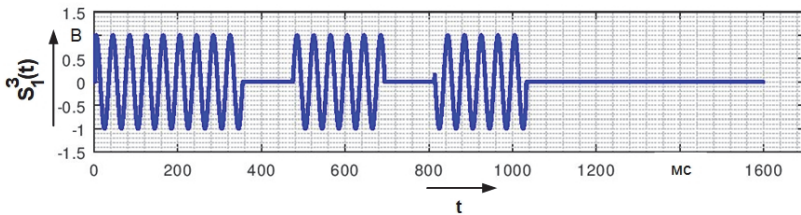


Рис. 6. График кодового сигнала 3

5. Моделирование искажений формы кодового сигнала.

Для моделирования искажений формы кодового сигнала примем, что они обусловлены: свойствами рельсовой линии, случайными внешними помехами, а также свойствами принимающей аппаратуры локомотива.

Согласно [20, 21] участок рельсовой линии заданной длины в электрических расчетах может быть представлен в виде четырехполюсника, схема замещения которого показана на рисунке 7а.

При этом экспериментальные исследования позволили установить, что емкостной составляющей $C_{из}$ при расчетах в случаях низко-

частотного несущего сигнала в рельсах (до 100 Гц) [20] можно пренебречь. Сопротивление рельсов R_p вносит пропорциональное своей величине затухание сигнала [22], что дает возможность им пренебречь при моделировании. Таким образом, схема на рисунке 7а может быть преобразована в схему, показанную на рисунке 7б. Схемы на рисунке 7 представляют собой фильтры нижних частот [23].

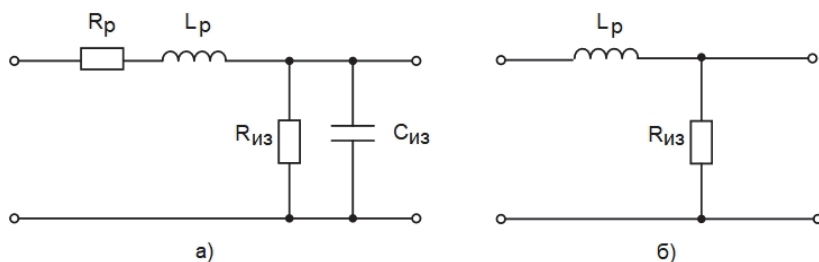


Рис. 7. Схема замещения четырехполюсника рельсовой линии: а – до преобразования, б – после преобразования (R_p – сопротивление рельсов (Ом/км), L_p – индуктивность рельсов (Гн/км), $R_{из}$ – сопротивление изоляции рельсовой линии (Ом·км), $C_{из}$ – емкость между рельсами (Ф/км))

Таким образом, в качестве модели рельсовой линии предлагается выбрать электрический фильтр нижних частот, частоту среза F_c которого можно определить по стандартной формуле (4) [24]:

$$F_c = \frac{R_{из}}{2\pi L_p}. \quad (4)$$

Сопротивление изоляции рельсовой линии $R_{из}$ и индуктивность рельсов L_p являются изменяемыми параметрами имитационной модели. При моделировании в качестве значений параметров $R_{из}$ и L_p приняты такие, согласно [21], при которых влияние рельсовой линии на кодовый сигнал максимально. Они составили 1 Ом·км и 2,3 мГн/км соответственно. Таким образом, частота среза F_c составила 68,9 Гц.

Для моделирования фильтра нижних частот в среде Octave использован цифровой фильтр Баттерворта нижних частот [23]. Выбор данного типа фильтра обусловлен «гладкостью» его амплитудно-частотной характеристики в полосе пропускания [23].

Передаточная функция цифрового фильтра описывается отношением (5) [25]:

$$H(z) = \frac{\sum_{i=0}^M b_i z^{-i}}{1 + \sum_{k=1}^N a_k z^{-k}}, \quad (5)$$

где M, N — количество нулей и полюсов цифрового фильтра соответственно; b_i, a_k — коэффициенты полиномов в числителе и знаменателе передаточной функции соответственно; z — оператор z -преобразования.

Коэффициенты передаточной функции цифрового фильтра Баттерворта нижних частот могут быть определены при помощи встроенной функции «butter» в среде Octave (6):

$$[b_{\text{пл}}, a_{\text{пл}}] = \text{butter}(n_{\text{пл}}, F_c), \quad (6)$$

где $b_{\text{пл}}, a_{\text{пл}}$ — коэффициенты передаточной функции цифрового фильтра Баттерворта нижних частот; $n_{\text{пл}}$ — порядок фильтра нижних частот, являющийся изменяемым параметром имитационной модели.

В результате расчета коэффициентов $b_{\text{пл}}$ и $a_{\text{пл}}$ согласно (6) передаточные функции выбранного цифрового фильтра имеют вид (7)-(10):

$$H(z)|_{n_{\text{пл}}=1} = \frac{0,18 + 0,18z^{-1}}{1 - 0,64z^{-1}}, \quad (7)$$

$$H(z)|_{n_{\text{пл}}=2} = \frac{0,04 + 0,07z^{-1} + 0,04z^{-2}}{1 - 1,40z^{-1} + 0,54z^{-2}}, \quad (8)$$

$$H(z)|_{n_{\text{пл}}=3} = \frac{10^{-1} \cdot (0,07 + 0,21z^{-1} + 0,21z^{-2} + 0,07z^{-3})}{1 - 2,14z^{-1} + 1,61z^{-2} - 0,42z^{-3}} \quad (9)$$

$$H(z)|_{n_{\text{пл}}=4} = \frac{10^{-2} \cdot (0,13 + 0,53z^{-1} + 0,79z^{-2} + 0,53z^{-3} + 0,13z^{-4})}{1 - 2,87z^{-1} + 3,21z^{-2} - 1,63z^{-3} + 0,31z^{-4}}. \quad (10)$$

Влияние рельсовой линии на кодовый сигнал может быть описано формулой (11):

$$S_1'(t) = \text{filter}(b_{\text{пл}}, a_{\text{пл}}, S_1(t)), \quad (11)$$

где «filter» — обозначение функции, моделирующей воздействие цифрового фильтра с заданными параметрами на сигнал в среде Octave.

Поскольку наибольшая вероятность возникновения внешних помех, приводящих к искажению формы кодового сигнала, наблюдается в рельсовой линии [2, 26], их моделирование предлагается реализовать в этом компоненте тракта. Воздействие помех на сигнал в рельсовой линии $S_1'(t)$ может быть записано в виде формулы (12):

$$S_2(t) = S_1'(t) + D(t). \quad (12)$$

Для моделирования данного воздействия на кодовый сигнал предлагается использовать Гауссовый «белый шум» (функция «awgn» в Octave). С учетом этого формула (12) будет иметь вид (13):

$$S_2(t) = \text{awgn}(S_1'(t), r), \quad (13)$$

где r — отношение «сигнал/шум», дБ.

Число r принято изменяемым параметром имитационной модели.

На рисунке 8 представлен график кодового сигнала $S_2^3(t)$ на выходе рельсовой линии с частотой несущего сигнала F_N , равной 25 Гц, с учетом внесенных в кодовый сигнал $S_1^3(t)$ искажений в виде «белого шума» по формуле (13). В качестве примера значение отношения «сигнал/шум» r принято равным 20 дБ.

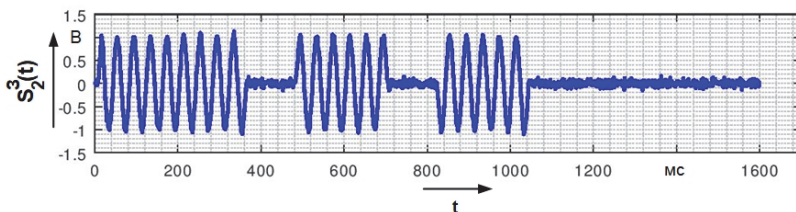


Рис. 8. График кодового сигнала 3 на выходе рельсовой линии

Локомотивный фильтр, входящий в состав бортового оборудования локомотива, предназначен для подавления внешних помех в принимаемом из рельсов кодовом сигнале [4]. Однако дополнительно он вносит и собственные искажения формы сигнала, что видно на осциллограмме реального сигнала на рисунке 2.

Согласно [5] локомотивный фильтр является полосовым. Для частот несущего сигнала 25 Гц и 50 Гц ширина полосы пропускания фильтра составляет 14 Гц. Номинальные значения граничных частот локомо-

тивного фильтра F_L (нижняя граница) и F_H (верхняя граница) при частотах несущего сигнала 25 Гц и 50 Гц представлены в таблице 3.

В качестве модели локомотивного фильтра выбран цифровой полосовой фильтр Баттерворта. При этом изменяемыми параметрами имитационной модели кодовых сигналов также будут:

- порядок полосового фильтра $n_{лф}$;
- ширина полосы пропускания полосового фильтра ΔF . Величину отклонения значений F_L и F_H от номинальных (таблица 3) предлагается варьировать опосредованно за счет изменения ширины полосы пропускания фильтра ΔF согласно формулам (14), (15).

Таблица 3. Номинальные значения граничных частот локомотивного фильтра при частотах несущего сигнала 25 Гц и 50 Гц

Частота несущего сигнала, Гц	F_L , Гц	F_H , Гц
25	18	32
50	43	57

В качестве модели локомотивного фильтра выбран цифровой полосовой фильтр Баттерворта. При этом изменяемыми параметрами имитационной модели кодовых сигналов также будут:

- порядок полосового фильтра $n_{лф}$;
- ширина полосы пропускания полосового фильтра ΔF . Величину отклонения значений F_L и F_H от номинальных (таблица 3) предлагается варьировать опосредованно за счет изменения ширины полосы пропускания фильтра ΔF согласно формулам (14), (15).

$$F_L = F_N - \Delta F/2, \quad (14)$$

$$F_H = F_N + \Delta F/2. \quad (15)$$

Таким образом, по аналогии с моделью рельсовой линии воздействие локомотивного фильтра на кодовый сигнал может быть описано формулами (16), (17):

$$[b_{лф}, a_{лф}] = \text{butter}(n_{лф}, [F_L, F_H]); \quad (16)$$

$$S_3(t) = \text{filter}(b_{лф}, a_{лф}, S_2(t)), \quad (17)$$

где $b_{лф}$, $a_{лф}$ — коэффициенты передаточной функции цифрового полосового фильтра Баттерворта.

Отметим, что в данном случае аргументами функции «butter» являются две частоты.

Передаточные функции данного цифрового фильтра, согласно (6), представлены соотношениями (18)-(21):

$$H(z)|_{n_{\text{лф}}=1} = \frac{0,04 - 0,04z^{-2}}{1 - 1,89z^{-1} + 0,92z^{-2}}; \quad (18)$$

$$H(z)|_{n_{\text{лф}}=2} = \frac{10^{-2} \cdot (0,18 + 0,36z^{-2} + 0,18z^{-4})}{1 - 3,83z^{-1} + 5,54z^{-2} - 3,60z^{-3} + 0,88z^{-4}}; \quad (19)$$

$$\begin{aligned} H(z)|_{n_{\text{лф}}=3} &= \\ &= \frac{10^{-3} \cdot (0,08 - 0,23z^{-2} + 0,23z^{-4} - 0,08z^{-6})}{1 - 5,75z^{-1} + 13,88z^{-2} - 17,93z^{-3} + 13,09z^{-4} - 5,12z^{-5} + 0,84z^{-6}}, \end{aligned} \quad (20)$$

$$\begin{aligned} H(z)|_{n_{\text{лф}}=4} &= \\ &= \frac{10^{-4} \cdot (0,03 - 0,13z^{-2} + 0,21z^{-4} - 0,13z^{-6} + 0,03z^{-8})}{1 - 7,68z^{-1} + 25,91z^{-2} - 50,09z^{-3} + 60,74z^{-4} - \\ &\quad - 47,29z^{-5} + 23,09z^{-6} - 6,47z^{-7} + 0,79z^{-8}}. \end{aligned} \quad (21)$$

В качестве примера на рисунке 9 представлен график сигнала $S_3^3(t)$ с F_N , равной 25 Гц, который получен по формулам (16), (17) при номинальных значениях F_L и F_H .

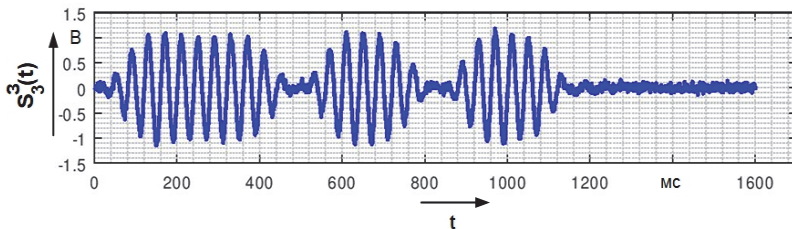


Рис. 9. График кодового сигнала 3 на выходе локомотивного фильтра

ГСА генерации сигнала $S_3^3(t)$, присутствующего на выходе локомотивного фильтра, представлена на рисунке 10. Для сигналов Ж и КЖ ГСА выглядит аналогично.

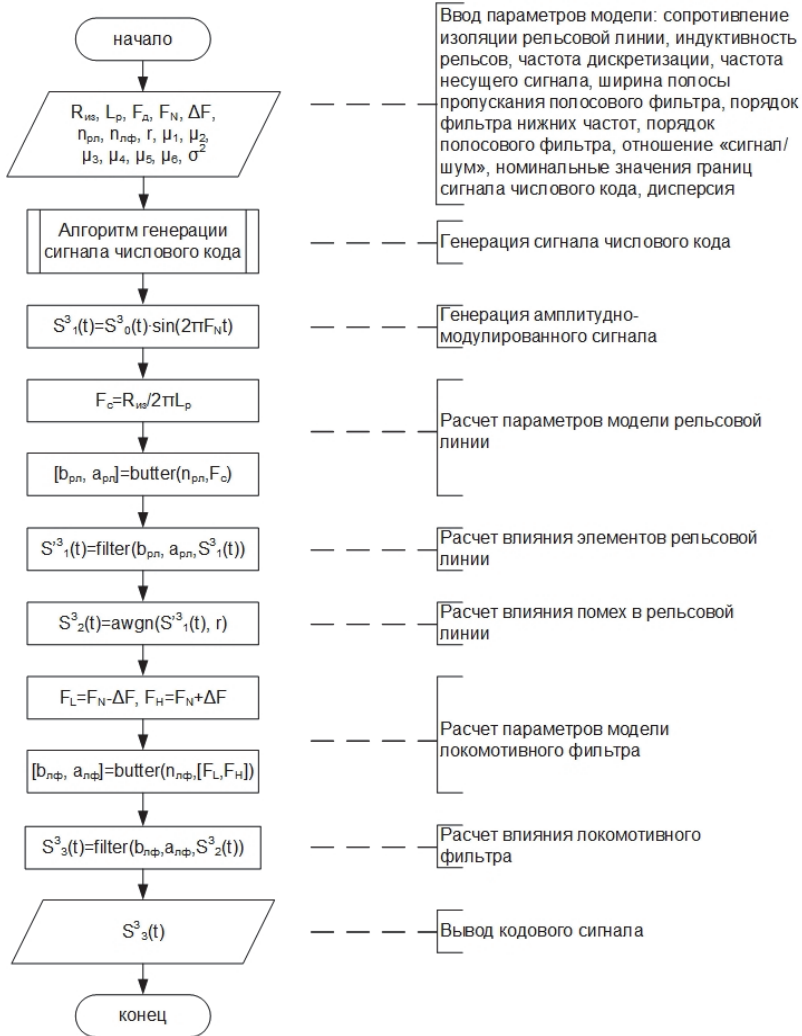


Рис. 10. ГСА генерации кодового сигнала $S^3_3(t)$, присутствующего на выходе локомотивного фильтра

6. Демонстрация имитационной модели. Таким образом, изменяемыми параметрами имитационной модели кодовых сигналов являются: сопротивление изоляции рельсовой линии; индуктивность рельсов; дисперсия для формирования случайного положения границы импульса или интервала внутри цикла сигнала числового кода;

порядок фильтра нижних частот для моделирования влияния рельсовой линии; отношение «сигнал/шум» при добавлении в кодовый сигнал «белого шума»; порядок полосового фильтра для моделирования влияния локомотивного фильтра; ширина полосы пропускания полосового фильтра, а также частота несущего сигнала. Диапазоны изменения указанных параметров имитационной модели приведены в таблице 4.

Таблица 4. Параметры имитационной модели и диапазоны их изменения

Параметр	Обозначение	Мин. значение	Макс. значение	Шаг
Сопrotивление изоляции рельсовой линии, Ом·км	$R_{из}$	1	—	—
Индуктивность рельсов, мГн/км	L_p	2,3	—	—
Дисперсия, ед.	σ^2	0,001	0,01	0,001
Порядок фильтра нижних частот, ед.	$n_{рл}$	1	4	1
Отношение «сигнал/шум», дБ	r	1	20	1
Порядок полосового фильтра, ед.	$n_{лф}$	1	4	1
Ширина полосы пропускания полосового фильтра, Гц	ΔF	10	14	1
Частота несущего сигнала, Гц	F_N	25	50	—

В результате генерации осциллограмм кодовых сигналов по представленной имитационной модели на всех возможных сочетаниях значений параметров (согласно таблице 4) получен массив данных размерности 96000×1600, соответствующий 96000 осциллограммам с 1600 отсчетами сигнала в каждой. Данный массив может быть использован для формирования обучающей выборки при обучении классификатора кодовых сигналов.

На рисунке 11 приведены девять демонстрационных осциллограмм кодовых сигналов, которые сгенерированы с помощью разрабо-

танной имитационной модели при различных сочетаниях значений ее параметров (таблица 5). Для наглядности на всех графиках частота несущего сигнала F_N равна 25 Гц.

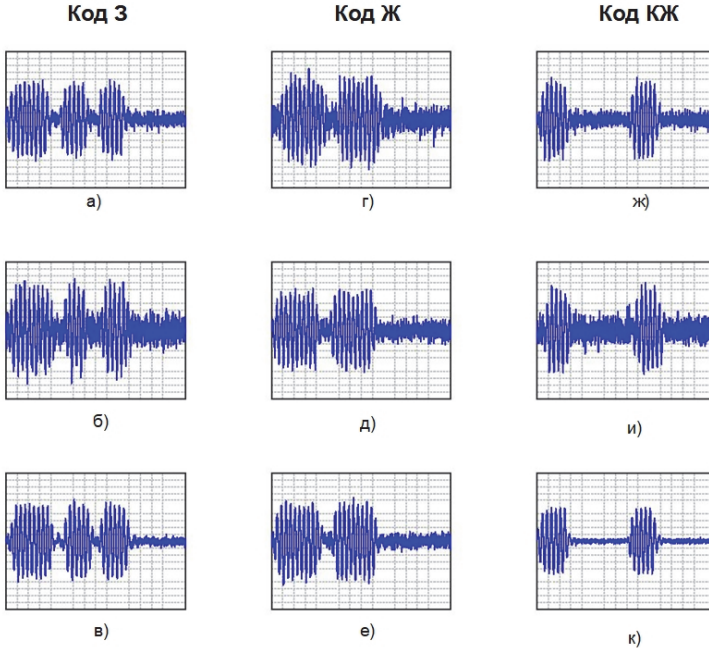


Рис. 11. Осциллограммы кодовых сигналов, сгенерированные при различных сочетаниях значений изменяемых параметров имитационной модели (а – в – код 3, г – е – код Ж, ж – к – код КЖ)

Таблица 5. Значения параметров имитационной модели для генерации демонстрационных осциллограмм кодовых сигналов

Обуч. пример	Кодовый сигнал	σ^2	$\mu_{рл}$	τ	$\mu_{лф}$	ΔF
а)	3	0,001	2	12	1	13
б)	3	0,008	1	6	2	12
в)	3	0,01	4	18	3	14
г)	Ж	0,001	3	8	4	10
д)	Ж	0,005	2	10	1	11
е)	Ж	0,01	3	13	2	13
ж)	КЖ	0,001	1	10	2	14
и)	КЖ	0,01	4	5	3	10
к)	КЖ	0,006	2	20	1	12

7. Оценка адекватности разработанной имитационной модели кодового сигнала. Для оценки адекватности разработанной имитационной модели кодового сигнала определим коэффициент корреляции реального сигнала, зафиксированного на борту локомотива (рисунок 1), $S_3^P(t)$ и смоделированного (рисунок 9) — $S_3^3(t)$ по формуле (22) согласно [17]:

$$K(S_3^P, S_3^3) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{S_{3i}^P - \mu_1}{\sigma_1} \right) \cdot \left(\frac{S_{3i}^3 - \mu_2}{\sigma_2} \right), \quad (22)$$

где $K(S_3^P(t), S_3^3(t))$ — коэффициент корреляции сигналов $S_3^P(t)$ и $S_3^3(t)$; N — число отсчетов (наблюдений) в сигналах $S_3^P(t)$ и $S_3^3(t)$ — равно 1600; μ_1, μ_2 — средние значения для $S_3^P(t)$ и $S_3^3(t)$ соответственно; σ_1, σ_2 — средние квадратические отклонения для $S_3^P(t)$ и $S_3^3(t)$ соответственно.

Коэффициент корреляции реального и синтезированного кодовых сигналов, полученный по формуле (22), составил 0,91. Таким образом, предлагаемую имитационную модель можно считать достаточно адекватной и пригодной для генерации обучающей выборки при построении классификатора кодовых сигналов.

8. Применение имитационной модели кодового сигнала для синтеза обучающей выборки. Для построения классификатора кодовых сигналов необходимо произвести предварительную обработку синтезированных имитационной моделью осциллограмм — сформировать обучающую выборку. Данная процедура сочетает в себе последовательное выполнение двух этапов:

- амплитудная демодуляция кодового сигнала;
- нормализация амплитудно-демодулированного сигнала.

Амплитудная демодуляция кодового сигнала позволяет выделить в его составе модулирующий сигнал, содержащий числовой код, по которому, в свою очередь, удобно выполнить классификацию.

Одним из вариантов реализации амплитудной демодуляции сигнала является синхронное детектирование [27]. Суть данного метода заключается в предварительном умножении исходного сигнала (в данном случае — кодового) на некоторый опорный сигнал, после чего в результате фильтрации происходит выделение модулирующего сигнала.

С учетом ранее введенных обозначений умножение синтезированного моделью кодового сигнала, в котором несущим, согласно (3), является сигнал $\sin(2\pi F_N t)$, на опорный сигнал может быть описано при помощи (23):

$$\begin{aligned}
 S_{\text{сд}}(t) &= S_3(t) \cdot \sin(2\pi F_N t) = S_{3A}(t) \cdot \sin^2(2\pi F_N t) = \\
 &= \frac{1}{2} S_{3A}(t) - \frac{1}{2} S_{3A}(t) \cdot \sin(4\pi F_N t),
 \end{aligned}
 \tag{23}$$

где $S_{\text{сд}}(t)$ — сигнал, полученный в результате умножения кодового сигнала $S_3(t)$ на несущий сигнал $\sin(2\pi F_N t)$; $S_{3A}(t)$ — изменяющаяся во времени амплитуда кодового сигнала $S_3(t)$.

Как следует из формулы (23), умножение кодового сигнала на опорный приводит к появлению суммы двух слагаемых. При этом первое слагаемое является искомым низкочастотным модулирующим сигналом. Для его выделения из $S_{\text{сд}}(t)$ в синхронном детектировании предусмотрена низкочастотная фильтрация, результатом которой является удаление из $S_{\text{сд}}(t)$ высокочастотной составляющей, то есть второго слагаемого. Низкочастотная фильтрация реализована путем применения к $S_{\text{сд}}(t)$ цифрового фильтра Баттерворта нижних частот, который уже был использован ранее. Порядок фильтра в данном случае, согласно [27], выбирается равным 5.

Результирующая амплитуда выделенного модулирующего сигнала по отношению к исходному, как следует из (23), уменьшается в два раза, что необходимо компенсировать. Такая компенсация реализована на этапе нормализации амплитудно-демодулированного сигнала.

Отметим, что выбор синусоидального сигнала с нулевой начальной фазой в качестве опорного обусловлен тем, что такой же сигнал использован в модели в качестве несущего на этапе модуляции (3). Это позволило обеспечить отсутствие фазовой ошибки, что необходимо при синхронном детектировании [27], поскольку ее наличие приводит к потере мощности сигнала.

Вторым этапом предварительной обработки кодового сигнала является нормализация выделенного на предыдущем этапе модулирующего сигнала. Под нормализацией в терминологии машинного обучения понимается представление значений обучающих примеров в унифицированном диапазоне, например от 0 до 1. Это способствует существенному повышению эффективности процедуры обучения классификатора за счет увеличения скорости обработки обучающих примеров [12].

Нормализация выделенного модулирующего сигнала в диапазоне от 0 до 1 может быть выполнена на основе (24):

$$x_i^{\text{норм}} = \frac{x_i - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})},
 \tag{24}$$

где \mathbf{x} — вектор, соответствующий кодовому сигналу после его амплитудной демодуляции (обучающий пример); i — число отсчетов в векторе \mathbf{x} (равно 1600); $x_i^{\text{норм}}$ — нормализованное значение отсчета в векторе \mathbf{x} ; $\min(\mathbf{x})$, $\max(\mathbf{x})$ — минимальное и максимальные значения отсчетов для вектора \mathbf{x} .

ГСА, поясняющая процесс предварительной обработки кодового сигнала, показана на рисунке 12.

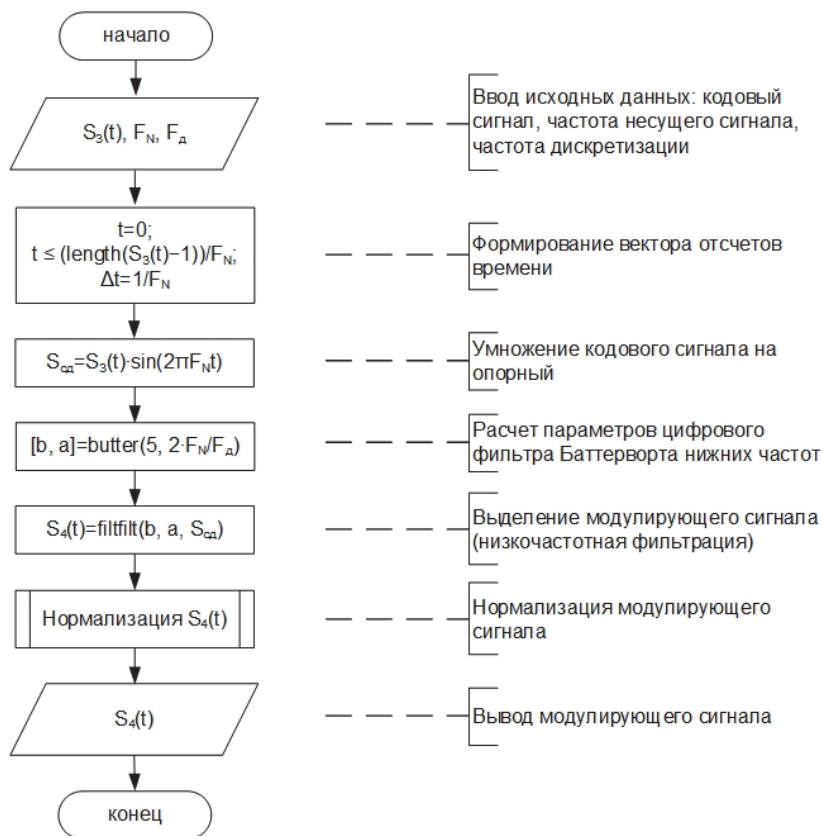


Рис. 12. ГСА предварительной обработки кодового сигнала

На рисунке 13 в качестве демонстрационных приведены обучающие примеры, полученные путем предварительной обработки осциллограмм кодовых сигналов на рисунке 11 а, д и к.

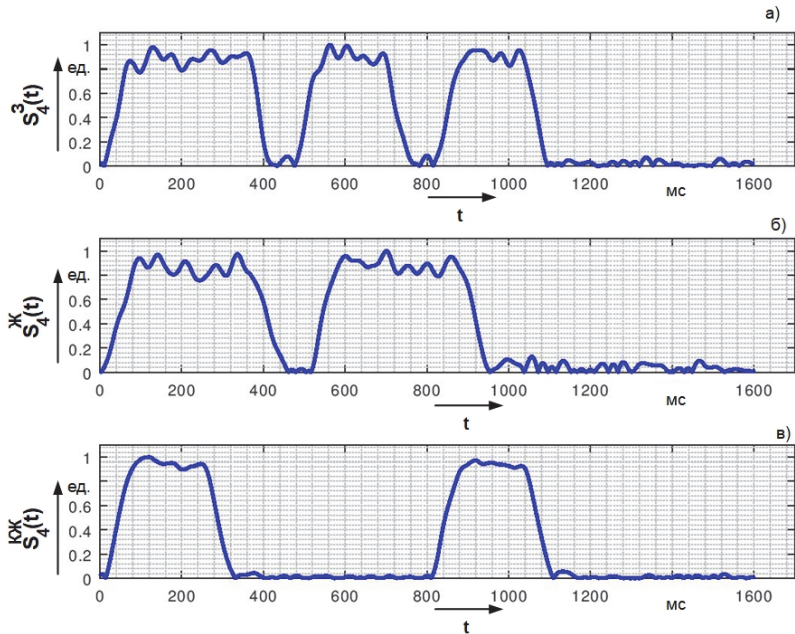


Рис. 13 Обучающие примеры, прошедшие предварительную обработку (а – для рисунка 11 а, б – для рисунка 11 д, в – для рисунка 11 к)

9. Заключение. Разработанная имитационная модель позволяет синтезировать выборку кодовых сигналов любого размера. Модель представляет собой алгоритм, описывающий основные этапы формирования кодовых сигналов. Благодаря модульному построению она может быть легко уточнена для учета процессов, свойств и явлений, которые не были приняты во внимание при разработке настоящего варианта модели.

Синтезированная выборка кодовых сигналов открывает возможность для разработки классификаторов на основе алгоритмов машинного обучения. Следует отметить, что наличие синтетической обучающей выборки не устраняет необходимости получения выборки реальных кодовых сигналов, а лишь позволяет без значительных экономических затрат проверить концепцию машинной классификации кодовых сигналов АЛСН. Очевидно, что оценить адекватность разработанной модели как источника обучающей выборки, можно путем разработки классификатора и проверки его работы на реальных кодовых сигналах.

Литература

1. Сапожников В.В. и др. Эксплуатационные основы автоматики и телемеханики // М.: Маршрут. 2006. 247 с.
2. Леушин В.Б., Юсупов Р.Р. Особенности каналов автоматической локомотивной сигнализации магистральных железных дорог // Самара: СамГУПС. 2007. 115 с.
3. Табуицков А.К., Горенбейн Е.В., Стрякин Л.И. Сбои АЛСН. Проблемы и пути их решения // Автоматика, связь, информатика. 2015. № 8. С. 21–22.
4. Брылеев А.М. и др. Автоматическая локомотивная сигнализация и авторегулировка // М.: Транспорт. 1981. 320 с.
5. Леонов А.А. Техническое обслуживание автоматической локомотивной сигнализации // М.: Транспорт. 1982. 255 с.
6. Шаманов В.И. Магнитные свойства рельсовых линий и уровень помех на аппаратуру автоматики и телемеханики // Электротехника. 2015. № 9. С. 50–54.
7. Лочехин В.С. Приемное устройство автоматической локомотивной сигнализации // Патент на изобретение № 2517631 РФ. Оpubл. 27.05.2014. Бюл. № 15.
8. Аргунов И.А. и др. Устройство подавления импульсных помех на входе локомотивного приемника АЛС // Патент на изобретение № 2618616 РФ. Оpubл. 04.05.2017. Бюл. № 13.
9. Леушин В.Б., Юсупов Р.Р., Блачев К.Э. Приемное устройство автоматической локомотивной сигнализации // Патент на полезную модель № 165420 РФ. Оpubл. 20.10.2016. Бюл. № 29.
10. Засов В.А., Железнов Д.В., Митрофанов А.Н., Белоногов А.С. Адаптивное подавление помех в приемных устройствах автоматической локомотивной сигнализации // Электротехника. 2017. № 3. С. 18–22.
11. Розенберг Е.Н. Цифровая железная дорога – ближайшее будущее // Автоматика, связь, информатика. 2016. № 10. С. 4–7.
12. Bishop C.M. Pattern recognition and machine learning // Springer. 2011. 738 p.
13. Goodfellow I., Bengio Y., Courville A. Deep Learning // MIT Press. 2016. 775 p.
14. Murphy K.P. Machine learning: a probabilistic perspective // MIT Press. 2012. 1104 p.
15. Harrington P. Machine Learning in Action // Manning Publications. 2012. 384 p.
16. Brink H., Richards J., Fetherolf M. Real-world machine learning // Manning Publications. 2016. 264 p.
17. Smith S.W. Digital signal processing: a practical guide for engineers and scientists // Newnes. 2013. 650 p.
18. Shai S. S., Shai B. D. Understanding machine learning: from theory to algorithms // Cambridge University Press. 2014. 410 p.
19. Сороко В.И., Фоткина Ж.В. Аппаратура железнодорожной автоматики и телемеханики // М.: НПФ «Планета». 2013. 1048 с.
20. Брылеев А.М., Крацов Ю.А., Шишляков А.В. Теория, устройство и работа рельсовых цепей // М.: Транспорт. 1978. 344 с.
21. Аркатов В.С., Аркатов Ю.В., Казеев С.В., Ободовский Ю.В. Рельсовые цепи магистральных железных дорог // ООО «Миссия-М». 2006. 496 с.
22. Alexander C.K., Sadiku M. Fundamentals of electric circuits // McGraw-Hill Education. 2016. 992 p.
23. Schubert T.F., Kim E.M. Fundamentals of electronics: book 3: active filters and amplifier frequency response // Morgan & Claypool Publishers. 2016. 294 p.
24. Bishop O. Understand Electronic Filters // Newnes. 1996. 180 p.
25. Lam H.Y-F. Analog and digital filters: design and realization // Prentice Hall. 1979. 632 p.
26. Пулятьков А.В., Скоробогатов М.Э. Системный анализ устойчивости работы систем автоматической локомотивной сигнализации // Современные технологии. Системный анализ. Моделирование. 2018. № 1. С. 79–89.
27. Сергиенко А.Б. Цифровая обработка сигналов // БХВ-Петербург. 2011. 768 с.

Присухина Илона Вадимовна — аспирант, кафедра автоматике и телемеханики, Омский государственный университет путей сообщения (ОмГУПС). Область научных интересов: теория и устройство рельсовых цепей, применение методов машинного обучения в системах железнодорожной автоматике и телемеханики. Число научных публикаций — 17. prisukhinaiv@omgups.ru; пр. Маркса, 35, 644046, Омск, Российская Федерация; р.т.: +7 (3812) 31-18-72.

Борисенко Дмитрий Владимирович — канд. техн. наук, доцент, кафедра автоматике и телемеханики, Омский государственный университет путей сообщения (ОмГУПС). Область научных интересов: машинное обучение, встраиваемые микропроцессорные системы, программное обеспечение микропроцессорных систем, моделирование микропроцессорных систем, безопасность систем железнодорожной автоматике и телемеханики. Число научных публикаций — 41. borisenkodv@yandex.ru; пр. Маркса, 35, 644046, Омск, Российская Федерация; р.т.: +7 (3812) 31-18-72.

Лунёв Сергей Александрович — канд. техн. наук, доцент, заведующий кафедрой, кафедра автоматике и телемеханики, Омский государственный университет путей сообщения (ОмГУПС). Область научных интересов: разработка средств измерения, контроля, технической диагностики устройств железнодорожной автоматике и телемеханики, безопасность систем железно-дорожной автоматике и телемеханики. Число научных публикаций — 116. lunevsa@omgups.ru; пр. Маркса, 35, 644046, Омск, Российская Федерация; р.т.: +7 (3812) 31-18-72.

I.V. PRISUKHINA, D.V. BORISENKO, S.A. LUNEV
**SIMULATION MODEL OF ELECTRIC CODE-MODULATED
SIGNAL IN RUSSIAN SYSTEMS OF INTERVAL CONTROL OF
TRAIN MOVEMENT BASED ON TRACK CIRCUIT**

Prisukhina I.V., Borisenko D.V., Lunev S.A. Simulation Model of Electric Code-Modulated Signal in Russian Systems of Interval Control of Train Movement Based on Track Circuit.

Abstract. Systems of interval control of train movement Signaling systems, which are currently in service in Russian railways, use the electric track circuit as the main data channel between signals and locomotives. Code-modulated electric signals transferred through that channel are frequently get corrupted which leads to railway traffic delays.

Decoding of the electric signal received from a track circuit can be represented as an image classification problem, and thus the stability of the data channel could be significantly improved.

However, to build such a classifier based on some machine learning algorithm, one needs a large dataset. In this article, a simulation model to synthesize this dataset is proposed.

The structure of the computer model matches the main stages of the electric code-modulated signal generation in a track circuit: code signal generator, rails, locomotive receiver.

Based on code signal generator schematic and waveform diagrams, a generator algorithm is developed. At this stage, we modeled timings of electric code signals according to the specification as well as their random deviations caused by various factors.

The analysis of substitution circuits of the rail line revealed that it has the properties of a low-pass filter. So, the rail line using the Butterworth digital filter with corresponding parameters is modeled. Additionally, at this stage, random noise during transmission was taken into account.

A similar technique is applied for modeling of a locomotive receiver which has a band-pass filter as the first signal processing block.

Thus, the proposed simulation model consists of a set of algorithms which run in series. By varying the parameters of the model, one can synthesize waveform diagrams of the electric code-modulated signal received by the locomotive equipment from a track circuit working in various modes and conditions.

Keywords: Code-Modulated Signal, Automatic Locomotive Signaling, Machine Learning, Butterworth Digital Filter, Rail Line, Normal Distributed Random Variable, Octave.

Prisukhina Iona Vadimovna — Ph.D. Student, Signaling and Interlocking Department, Omsk State Transport University (OSTU). Research interests: track circuits basics and design, machine learning algorithms applications in signaling and interlocking systems. The number of publications — 17. prisukhinaiv@omgups.ru; 35, Marx pr., 644046, Omsk, Russian Federation; office phone: +7 (3812) 31-18-72.

Borisenko Dmitry Vladimirovich — Ph.D., Associate Professor, Signaling and Interlocking Department, Omsk State Transport University (OSTU). Research interests: machine learning, embedded microprocessor systems, microprocessor systems soft-ware, microprocessor systems modeling, signaling and interlocking systems safety. The number of publications — 41. borisenkodv@yandex.ru; 35, Marx pr., 644046, Omsk, Russian Federation; office phone: +7 (3812) 31-18-72.

Lunev Sergey Aleksandrovich — Ph.D., Associate Professor, Head of Department, Signaling and Interlocking Department, Omsk State Transport University (OSTU). Research interests: signaling and interlocking systems measurement, control and diagnostics tool design, signaling and interlocking systems safety. The number of publications — 116. lunevsa@omgups.ru; 35, Marx pr., 644046, Omsk, Russian Federation; office phone: +7 (3812) 31-18-72.

References

1. Sapozhnikov V.V. et al. *Ehkspluatatsionnye osnovy avtomatiki i telemekhaniki* [Operational basics of automation and remote control]. M.: Marshrut Publ. 2006. 247 p. (In Russ).
2. Leushin V.B., Yusupov R.R. *Osobennosti kanalov avtomaticheskoy lokomotivnoy signalizatsii magistral'nykh zheleznnykh dorog* [Features of channels of automatic locomotive signaling of main railways]. Samara State Transport University. 2007. 115 p. (In Russ).
3. Tabunshchikov AK, Gorenbeyn E.V., Stryapkin L.I. [Failures of automatic locomotive signaling. Problems and solutions]. *Avtomatika, svyaz', informatika – Automation, communication and Informatics*. 2015. vol. 8. pp. 21–22. (In Russ).
4. Bryleev A.M., Pope O., Dmitriev V.S., Kravtsov Y.A., Stepensky B.M. *Avtomaticheskaya lokomotivnaya signalizatsiya i avtoregulirovka* [Automatic locomotive signaling and auto-adjustment]. M.: Transport Publ. 1981. 320 p. (In Russ).
5. Leonov A.A. *Tekhnicheskoe obsluzhivanie avtomaticheskoy lokomotivnoy signalizatsii* [Maintenance of automatic locomotive signaling]. M.: Transport Publ. 1982. 255 p. (In Russ).
6. Shamanov V.I. [The magnetic properties of rail lines and level of interferences for the apparatus of automatic control and telemechanics]. *Elektrotekhnika – Russian Electrical Engineering*. 2015. vol. 9. pp. 50–54. (In Russ).
7. Lochehin V.S. [Receiving device for automatic cab signalling]. *Invention № 2517631 RU*. Publ. 27.05.2014. Paper № 15. (In Russ).
8. Argunov I.A. et al. [Device for suppressing impulse noise in input of locomotive receiver ALS]. *Invention № 2618616 RU*. Publ. 04.05.2017. Paper № 13. (In Russ).
9. Leushin V.B., Yusupov R.R., Blachev K.E. [Receiving device for automatic cab signalling]. *Utility model № 165420 RU*. Publ. 20.10.2016. Paper № 29. (In Russ).
10. Zasov V.A., Zheleznov D.V., Mitrofanov A.N., Belonogov A.S. [Adaptive noise abatement in automatic cab signaling receivers]. *Elektrotekhnika – Russian Electrical Engineering*. 2017. vol. 3. pp. 18–22. (In Russ).
11. Rosenberg E.N. [Digital Railway - Nearest Future]. *Avtomatika, svyaz', informatika – Automation, communication and Informatics*. 2016. vol. 10. pp. 4–7. (In Russ).
12. Bishop C.M. *Pattern recognition and machine learning*. Springer. 2011. 738 p.
13. Goodfellow I., Bengio Y., Courville A. *Deep Learning*. MIT Press. 2016. 775 p.
14. Murphy K.P. *Machine learning: a probabilistic perspective*. MIT Press. 2012. 1104 p.
15. Harrington P. *Machine Learning in Action*. Manning Publications. 2012. 384 p.
16. Brink H., Richards J., Fetherolf M. *Real-world machine learning*. Manning Publications. 2016. 264 p.
17. Smith S.W. *Digital signal processing: a practical guide for engineers and scientists*. Newnes. 2013. 650 p.
18. Shai S.S., Shai B.D. *Understanding machine learning: from theory to algorithms*. Cambridge University Press. 2014. 410 p.
19. Soroko VI, Fotkina Z.V. *Apparatura zheleznodorozhnoy avtomatiki i telemekhaniki* [Railway automatics and remote control equipment]. Planeta Publ. 2013. 1048 p. (In Russ).
20. Bryleev A.M., Kravtsov Y.A., Shishlyakov A.V. *Teoriya, ustrojstvo i rabota rel'sovykh cepej* [Theory, design and operation of track circuits]. M.: Transport Publ. 1978. 344 p. (In Russ).
21. Arkatov V.S., Arkatov Y.V., Kazeev S.V., Obodovskiy Y.V. *Rel'sovye cepi magistral'nykh zheleznnykh dorog* [Track circuits of main railways]. M.: Missia-M Publ. 2006. 496 p. (In Russ).

22. Alexander C.K., Sadiku M. Fundamentals of electric circuits. McGraw-Hill Education. 2016. 992 p.
23. Schubert T.F., Kim E.M. Fundamentals of electronics: book 3: active filters and amplifier frequency response. Morgan & Claypool Publishers. 2016. 294 p.
24. Bishop O. Understand Electronic Filters. Newnes. 1996. 180 p.
25. Lam H.Y-F. Analog and digital filters: design and realization. Prentice Hall. 1979. 632 p.
26. Pulyakov A.V, Skorobogatov M.E. [System analysis of sustainability of automatic locomotive signaling systems]. *Sovremennye tekhnologii. Sistemnyj analiz. Modelirovanie – Modern technologies. System analysis. Modeling.* 2018. vol. 1. pp. 79–89. (In Russ).
27. Sergienko A.B. *Cifrovaya obrabotka signalov* [Digital Signal Processing]. BHV-Petersburg Publ. 2011. 768 p. (In Russ).

В.И. МИРОНОВ, Ю.В. МИРОНОВ, Д.К. ХЕГАЙ
**ОПТИМАЛЬНОЕ ОПРЕДЕЛЕНИЕ ОРБИТЫ КОСМИЧЕСКИХ
ОБЪЕКТОВ ПО УГЛОВЫМ ИЗМЕРЕНИЯМ НАЗЕМНЫХ
ОПТИКО-ЭЛЕКТРОННЫХ СТАНЦИЙ**

Мионов В.И., Мионов Ю.В., Хегай Д.К. Оптимальное определение орбиты космических объектов по угловым измерениям наземных оптико-электронных станций.

Аннотация. Успешное решение задач практической космонавтики во многом обеспечивается современными достижениями в области измерительной и вычислительной техники, а также совершенством методов первичной и вторичной обработки траекторных измерений. Поэтому в перспективных программах освоения космического пространства и развития космической техники большое внимание уделяется совершенствованию существующих и разработки новых алгоритмических и технических средств навигационного обеспечения полетов космических объектов в интересах расширения возможностей и повышения эффективности систем автономной навигации космических аппаратов, а также наземных и перспективных орбитальных систем контроля космического пространства. В настоящее время ведется активная работа по модернизации и развитию перспективных комплексов специализированных оптико-электронных средств для мониторинга околоземного космического пространства на основе проводимых угловых измерений. Рассматривается применение вариационного подхода для решения задач статистического оценивания параметров траектории движения орбитального объекта по результатам угловых измерений, проводимых наземными оптико-электронными средствами, которые входят в состав современной системы контроля космического пространства. Приводятся модели и алгоритмы определения оценок параметров орбиты, реализующие вариационный вариант метода максимального правдоподобия, а также результаты тестовых расчетов, связанные с итерационным решением двухточечной краевой задачи вариационного оценивания.

Основная цель численных расчетов — исследование сходимости предлагаемого алгоритма оценивания, а также влияния ошибок измерений на смещение получаемых оценок относительно их точных значений.

Приведенные в статье результаты моделирования соответствуют условиям орбитального движения космического аппарата METEOR PRIRODA и получены с использованием эфемеридных данных каталога NORAD в TLE-элементах.

Ключевые слова: статистическое оценивание, космический аппарат, нелинейные динамические системы, критерий максимального правдоподобия, оптические угловые измерения.

1. Введение. Задачи оценивания параметров орбитального движения космических объектов (КО) по результатам измерений широко применяются на практике. Наиболее сложные задачи оценивания приходится решать, в частности, при навигационно-баллистическом обеспечении полетов космических аппаратов (КА), при разработке систем автономной навигации, а также в наземных и перспективных орбитальных системах контроля космического пространства.

В силу указанных причин в работах последних лет вопросам совершенствования методов определения орбит КО продолжает уде-

латься большое внимание. Принципиально для решения этих задач могут применяться как методы обработки навигационных измерений по полной выборке [1-5], так и методы динамической фильтрации [3], основанные на применении известного фильтра Калмана и его модификаций. Возможности и целесообразность применения того или иного подхода во многом определяется спецификой решаемых задач, составом измерений, а также требованиями, предъявляемыми к точности, надежности и оперативности решения соответствующих навигационных задач.

Применительно к задачам определения орбит КО по результатам угловых измерений наземных оптико-электронных станций (ОЭС), как показывают результаты исследований и накопленный опыт, использование алгоритмов традиционного линеаризованного фильтра Калмана не обеспечивает требуемой эффективности. Это объясняется существенной нелинейностью задачи оценивания, обусловленной нелинейностью динамической модели объекта наблюдения и модели угловых измерений, которая приводит к изменению закона распределения ошибок при дискретной обработке измерений, и влиянием других факторов [3, 5]. Отмечено, что в рассматриваемом классе задач определения орбит лежащие в основе рекуррентных алгоритмов байесовские условные функции плотности вероятностей могут существенно отличаться от гауссовской. Указанные факторы часто приводят к повышенным ошибкам определения параметров орбит и даже к расходимости вычислительного процесса оценивания.

В последние годы вопросы разработки новых типов нелинейных фильтров Калмана для определения орбит КО по результатам обработки угловых измерений рассматривались, в частности, в работах [6-13]. Основное направление этих работ состоит в стремлении учесть негауссовский характер этой проблемы при помощи плотности вероятностей функций, которые являются более общими, чем распределение Гаусса.

Так, в [9] для обеспечения точности и надежности динамической фильтрации угловых измерений предложено использовать распределение Мизеса. В работе [10] для решения задач нелинейной фильтрации предлагается использовать модель полиномиального хаоса. Однако осталось неясным, как можно ввести полиномы хаоса в рекурсивный алгоритм оценивания. В [7, 8, 11-13] исследована возможность применения Гауссовых смесей различной степени сложности. В работах [11, 12] предложены специальные процедуры, позволяющие контролировать точность аппроксимации условной байесовской плотности вероятности. Этот методический подход получил дальнейшее развитие в [6] и [13] в части обобщения результатов, полученных

в [11] и [12]. К указанному направлению исследований относится и работа [14], в которой используются дополнительно мягкие ограничения на допустимую область при формировании Гауссовых смесей.

Предложенные в рассмотренных выше работах методические средства расширяют возможности динамической фильтрации. Однако они не позволяют во всех случаях получать надежный результат и не решают проблему точности и надежности навигационных определений применительно к рассматриваемому классу задач оценивания в целом [3, 5].

Ряд исследований, проведенных в последние годы, посвящены проблематике создания орбитальных систем оптического наблюдения космического пространства в интересах обеспечения астероидной и кометной безопасности Земли, а также развитию наземной сети ОЭС станций для определения орбит элементов космического мусора и прогнозирования опасных ситуаций, связанных с возможным столкновением с действующими КА [5, 18-20].

Особую значимость в решении проблем построения системы контроля орбит космического мусора имеет работа [5], в которой рассмотрен весь комплекс вопросов, связанных с состоянием и перспективами развития наземной сети ОЭС, организации и методического обеспечения процессов определения орбит элементов космического мусора в околоземном космическом пространстве. При этом особо отмечается роль методов статистической обработки по полной выборке измерений в рассматриваемых задачах определения орбит.

В [18] представлена концепция космической системы обнаружения опасных небесных тел. Предложен вариант компоновки КА с использованием малоразмерной платформы, а также даны предварительные оценки основных параметров системы. В [19] исследованы различные варианты организации космического патруля для обнаружения и каталогизации опасных для Земли КО. В их основе лежит идея создания «оптического барьера» с помощью телескопов, размещенных на гелиоцентрической орбите. В [20] проведен анализ характеристик точности определения орбит, рассчитанных по коротким дугам наблюдений. При этом ставилось условие, что длина дуги наблюдений должна обеспечить уверенную классификацию орбит малых небесных тел, которая позволила бы выделить потенциально опасные тела, угрожающие столкновением.

Работа [21] посвящена анализу влияния различных возмущающих факторов на высокоточный прогноз орбит КО. Проведена систематизация возмущений для высокоорбитальных и низкоорбитальных КО. Дана оценка их степени влияния и разработаны модели возмуще-

ний, обеспечивающие существенное повышение точности прогноза эфемерид. Полученные результаты имеют немаловажное значение для обеспечения точности навигационных определений траекторий движения орбитальных объектов, а также для надежного прогнозирования возможных столкновений функционирующих КА с элементами космического мусора.

На основании проведенного выше анализа можно сделать вывод, что в настоящее время основными методами определения орбит КА являются методы, реализующие совместную обработку результатов наблюдений по полной выборке. Они широко освещены в отечественной и зарубежной литературе и успешно решают широкий круг важных и сложных прикладных задач. Однако вопросы улучшения их точностных и вычислительных характеристик продолжают оставаться актуальными.

Созданная методология базируется на непосредственном применении в динамических задачах оценивания условий метода максимального правдоподобия (ММП) и метода наименьших квадратов. По смыслу они представляют собой необходимые условия оптимальности, характерные для прямых методов оптимизации.

Вместе с тем необходимо отметить, что методы теории оптимальной обработки измерений, как и методы теории оптимального управления, могут строиться и развиваться на основе использования различных форм и принципов формирования условий оптимальности — как прямых, так и вариационных. Вариационные условия оптимальности создают новую базу для решения данного класса задач.

Каждая форма условий оптимальности имеет свою область наиболее рационального применения. Теоретический анализ и накопленный опыт практического применения методов теории оптимального управления убедительно показывает, что там, где удастся реализовать вариационные условия оптимальности (принцип максимума), обеспечиваются лучшие точностные и вычислительные характеристики алгоритмов по сравнению с алгоритмами управления, основанными на прямых методах оптимизации. Поэтому можно ожидать, что разработка и применение вариационного подхода к задачам определения орбит КО по угловым измерениям ОЭС также позволит обеспечить необходимые вычислительные и точностные характеристики соответствующих алгоритмов.

Вопросы теоретического обоснования и применения указанного вариационного подхода к задачам статистического оценивания нелинейных динамических систем рассматривались в работах авторов [22, 23].

Данная статья посвящена вопросам применения указанного вариационного подхода к решению задач оценивания параметров траек-

тории движения орбитального объекта по результатам угловых измерений, проводимых наземными ОЭС, с использованием критерия максимального правдоподобия. Приводятся модели динамики КО, модели измерений, соответствующие системы сопряженных дифференциальных уравнений краевой задачи, алгоритм определения оценок параметров орбиты, а также результаты тестовых расчетов применительно к условиям полета КО МЕТЕОР – ПРИРОДА (по международной классификации — METEOR PRIRODA).

Характеристики сходимости итерационных процессов вариационного оценивания параметров орбиты КО во многом зависят от точности определения начального приближения неизвестного вектора кинематических параметров. Для получения практически приемлемых результатов здесь могут быть использованы приближенные численно-аналитические решения, которые приведены в работах [4, 24, 25], полученные применительно к условиям движения КО в центральном гравитационном поле Земли по трем или четырем измерительным точкам. Этот вопрос кратко рассматривается ниже в пункте 5.

2. Общая постановка задачи и вариационные условия оптимальности оценок. Рассмотрим задачу оценивания параметров движения динамического объекта, которая заключается в наилучшем в некотором смысле определении n -мерного вектора его исходного состояния \bar{x}_0 на заданный начальный момент времени $t = t_0$ по результатам измерений, проводимых в N точках t_i , заданных на интервале измерений $\tau = T - t_0$, где T — конечный момент интервала измерений.

Пусть динамика объекта описывается векторным дифференциальным уравнением:

$$\dot{\bar{x}} = \bar{\varphi}(\bar{x}, t), \quad \bar{x}(t_0) = \bar{x}_0, \quad t \in [t_0, T]. \quad (1)$$

Измерениям подвергается m -мерный вектор:

$$\bar{\psi}(t) = \bar{\psi}[\bar{q}(t)].$$

Измеренное значение вектора $\bar{\psi}$ в момент t_i обозначим как $\bar{y}(t_i) = \bar{y}_i$, и представим модель измерений в виде:

$$\begin{aligned} \bar{y}(t_i) &= \bar{\psi}[\bar{x}(t_i)] + \bar{\delta}_i; \\ i &= 1(1)N; \quad t_i \in [t_0, T]. \end{aligned}$$

Здесь $\bar{\delta}_i$ — m -мерный вектор случайных ошибок измерений, стохастическое изменение которого зададим некоторым многомерным непрерывным дифференцируемым распределением $f(\bar{\delta}_i, \bar{\alpha}_i)$ с параметрами $\bar{\alpha}_i$, отличающемся в общем случае от нормального распределения.

Требуется найти такую оценку вектора \bar{x}_0 , которая обеспечивает экстремальное значение функционала:

$$I = \sum_{i=1}^N \rho_i \{ \bar{y}(t_i), \bar{\psi}[\bar{x}(t_i), \bar{c}_1], \bar{\alpha}_i \}, \quad (2)$$

где

$$\rho_i = \ln f_i \{ \bar{y}(t_i) - \bar{\psi}[\bar{x}(t_i), \bar{c}_1], \bar{\alpha}_i \}; \quad i = 1(1)N.$$

Данный функционал является логарифмической функцией правдоподобия. Предполагается выполнение известных условий наблюдаемости.

В работе [22] показано, что задача определения оптимальной оценки вектора \bar{x}_0 и порождаемой ей оптимальной траектории $\bar{x}(\bar{x}_0, t)$ сводится к решению двухточечной краевой задачи для канонической системы:

$$\dot{\bar{x}} = \frac{\partial H}{\partial \bar{x}}; \quad \dot{\bar{\lambda}} = -\frac{\partial H}{\partial \bar{x}},$$

где $H = \bar{\lambda}^T \bar{\varphi}(\bar{x}, t)$ — гамильтониан системы; $\bar{\lambda}$ — n -мерная вектор-функция сопряженных переменных с промежуточными и граничными (краевыми) условиями:

$$\begin{aligned} \bar{\lambda}(t_i^+) &= \bar{\lambda}(t_i^-) + \frac{\partial \rho}{\partial \bar{x}}[\bar{y}_i, \bar{\psi}(\bar{x}_i), t_i]; \\ \bar{\lambda}(t_0) &= 0; \quad \bar{\lambda}(T) = 0. \end{aligned}$$

В совокупности эти соотношения характеризуют вариационные условия оптимальности оценок. Таким образом, рассматриваемую задачу можно интерпретировать как двухточечную краевую задачу с промежуточными ограничениями на сопряженный вектор $\bar{\lambda}$.

При нормальном законе распределения ошибок измерений $N(0, K_{\delta_i})$ с нулевым вектором математического ожидания и корреляционной матрицей K_{δ_i} необходимо найти такую оценку вектора \bar{x}_0 , которая обеспечивает минимальное значение функционала:

$$J = \sum_{i=1}^N \{ \bar{y}(t_i) - \bar{\psi}[\bar{x}(t_i)] \}^T K_{\delta_i}^{-1} \{ \bar{y}(t_i) - \bar{\psi}[\bar{x}(t_i)] \}. \quad (3)$$

В этом случае краевая задача вариационного оценивания представляется следующими уравнениями и условиями:

$$\dot{\bar{x}} = \frac{\partial H}{\partial \bar{x}}; \quad \dot{\bar{\lambda}} = -\frac{\partial H}{\partial \bar{x}}; \quad (4)$$

$$\begin{aligned} \bar{\lambda}(t_i^+) &= \bar{\lambda}(t_i^-) + \frac{\partial \bar{\psi}_i^T}{\partial \bar{x}_i} K_{\delta_i}^{-1} \{ \bar{y}_i - \bar{\psi}[\bar{x}(t_i)] \}; \quad i = 1(1)N; \\ \bar{\lambda}(t_0) &= 0; \quad \bar{\lambda}(T) = 0. \end{aligned} \quad (5)$$

Для решения данной краевой задачи заметим, что оптимальная оценка вектора параметров движения объекта \bar{x}_0 является корнем уравнения:

$$\bar{\lambda}(\bar{x}_0, T) = 0, \quad (6)$$

заданного неявно на процедурах интегрирования системы (4).

Здесь $\bar{\lambda}(\bar{x}_0, T)$ есть значение сопряженного вектора, определенное для конечного момента времени $t = T$ - мерного интервала.

Для решения указанного уравнения могут быть использованы различные численные методы, такие как метод Ньютона и другие. Вычислительная схема метода Ньютона имеет вид:

$$\begin{aligned} \bar{x}_{0_{i+1}} &= \bar{x}_{0_i} - \left[\frac{\partial \bar{\lambda}(\bar{x}_0, T)}{\partial \bar{x}_0} \right]_i^{-1} \bar{\lambda}(\bar{x}_{0_i}, T); \\ \det \left[\frac{\partial \bar{\lambda}(\bar{x}_0, T)}{\partial \bar{x}_{0_i}} \right] &\neq 0, \end{aligned} \quad (7)$$

где i — номер итерации.

Для вычисления матрицы частных производных, входящей в (7), можно использовать численный метод конечных разностей.

Перейдем к непосредственному рассмотрению вопросов применения данного вариационного подхода к решению задач определения параметров траектории орбитального объекта по результатам угловых измерений, проводимых наземными ОЭС.

3. Математические модели краевой задачи вариационного определения параметров траектории орбитального объекта по результатам угловых измерений, проводимых наземными оптоэлектронными средствами. Для реализации изложенного выше вариационного подхода к определению параметров траектории орбитального объекта по результатам угловых измерений, проводимых наземными ОЭС, необходимо, прежде всего, конкретизировать основные математические модели, входящие в постановку краевой задачи оценивания. Поэтому рассмотрим уравнения движения КО, математические модели измерений, а также сопряженную систему дифференциальных уравнений.

Уравнения движения КО. При решении задач определения параметров траектории орбитального объекта по результатам угловых измерений, проводимых наземными ОЭС, будем использовать относительную гринвичскую систему координат (ГСК) $Oxyz$ с началом в центре масс Земли, плоскость (xy) совпадает с плоскостью экватора, ось x направлена в нулевой (Гринвичский) меридиан, ось z — по вектору угловой скорости вращения Земли, а ось y дополняет систему до правой.

Уравнения движения КО могут быть представлены с любой требуемой степенью сложности в зависимости от применяемых моделей действующих сил. Это не является ограничивающим фактором для применения вариационного варианта ММП для статистической обработки измерений. Если ограничиться учетом второй зональной гармоники в разложении геопотенциала, то уравнения движения КО можно представить в следующем удобном для программирования виде [1]:

$$\begin{aligned} \dot{x} &= V_x; & \dot{y} &= V_y; & \dot{z} &= V_z; \\ \dot{V}_x &= (\Omega_3^2 - A)x + 2\Omega_3 V_y - SpVV_x; \\ \dot{V}_y &= (\Omega_3^2 - A)y + 2\Omega_3 V_x - SpVV_y; \\ \dot{V}_z &= (2BC - A)z - SpVV_z, \end{aligned} \tag{8}$$

где

$$\begin{aligned}
 A &= B[\alpha_{00} + C(D-1)]; \quad B = \frac{1}{r^2} \frac{R_{\text{Э}}}{r}; \\
 C &= \frac{3}{2} \alpha_{20} \left(\frac{R_{\text{Э}}}{r} \right)^2; \quad D = 5 \left(\frac{z}{r} \right)^2; \\
 r &= \sqrt{x^2 + y^2 + z^2}; \quad r_1 = \sqrt{x^2 + y^2}; \\
 V &= \sqrt{V_x^2 + V_y^2 + V_z^2}; \quad \alpha_{00} = \frac{\mu}{R_{\text{Э}}}; \quad \alpha_{20} = -\frac{c_{20}}{R_{\text{Э}}^3};
 \end{aligned}$$

x, y, z, V_x, V_y, V_z — составляющие векторов координат и скорости КА в ГСК; S — баллистический коэффициент; ρ — плотность воздуха; μ, c_{20} — параметры модели ГПЗ; $R_{\text{Э}}$ — экваториальный радиус Земли.

При расчете силы сопротивления воздуха принимается, что атмосфера вращается вместе с Землей. Влияние ветра на движение КА не учитывается. При таких допущениях $V_{\text{отн}} = V$.

Модель измерений. Модели измерений представляют собой математические зависимости между измеряемыми и текущими кинематическими параметрами движения КО, а также ошибками измерений.

Независимым измерениям с помощью ОЭС наземного измерительного пункта (ИП) подвергается двумерный вектор:

$$\bar{\psi}(t) = \bar{\psi}[\bar{q}(t)], \quad (9)$$

где

$$\begin{aligned}
 \bar{\psi}(t) &= [A(t), \gamma(t)]^T; \quad \bar{q}(t) = [\bar{r}(t), \bar{V}(t)]^T; \\
 \bar{r}(t) &= [x(t), y(t), z(t)]^T; \quad \bar{V}(t) = [V_x(t), V_y(t), V_z(t)]^T;
 \end{aligned} \quad (10)$$

$A(t)$ — азимут объекта; $\gamma(t)$ — угол места объекта; $\bar{r}(t)$, $\bar{V}(t)$ — векторы координат и скорости КО в ГСК соответственно.

Измерительные ОЭС имеют соответствующие географические координаты точек стояние в ГСК: широту $B_{\text{П}}$, долготу $L_{\text{П}}$ и высоту $h_{\text{П}}$.

Измеренные значение вектора $\bar{\psi}_i$ в моменты t_i обозначим через \bar{y}_i и представим модель измерений в виде:

$$\begin{aligned} \bar{y}(t_i) &= \bar{\psi}[\bar{q}(t_i)] + \bar{\delta}_i; \\ i &= 1(1)N; \quad t_i \in [t_1, t_N], \end{aligned} \quad (11)$$

где N — число измерительных точек; t_1 — момент первого измерения, $t_1 \geq t_0$; t_N — момент последнего измерения, $t_N > T$.

$\bar{\delta}_i$ — m -мерный вектор случайных ошибок измерений.

Будем полагать, что $\bar{\delta}_i$ подчиняется нормальному закону распределения с нулевым математическим ожиданием и корреляционной матрицей K_{δ_i} :

$$\delta_i \rightarrow N(0, K_{\delta_i}),$$

где

$$K_{\delta_i} = \begin{bmatrix} \sigma_{A_i}^2 & K_i \\ K_i & \sigma_{\gamma_i}^2 \end{bmatrix}.$$

Здесь σ_{A_i} и σ_{γ_i} — среднеквадратические ошибки измерений азимута и угла места КО в соответствующие моменты измерений t_i , $i = 1(1)N$, K_i — корреляционные моменты ошибок измерений.

Приведем расчетные соотношения, определяющую связь измеряемых угловых параметров A , γ с кинематическими параметрами траектории КА и координатами ОЭС, заданными в ГСК.

Обозначим через $\bar{\rho} = [\xi, \eta, \zeta]^T$ вектор координат КО, определяемый в топоцентрической измерительной системе координат (ИСК) $O_{II}\xi, \eta, \zeta$. В этой системе точка O_{II} находится в точке расположения измерительного пункта, ось $O_{II}\xi$ направлена на северный полюс Земли по касательной к меридиану точки O_{II} , ось $O_{II}\eta$ — по внешней нормали к земному эллипсоиду, а ось $O_{II}\zeta$ дополняет систему до правой. Тогда расчетные значения измеряемых параметров $\bar{\psi}_i(t) = \bar{\psi}[\bar{q}(t_i)]$, входящие в оптимизируемый функционал и в формулы определения скачков сопряженных переменных (5), будут вычисляться путем интегрирования уравнений движения в ГСК на моменты t_i по следующим формулам:

$$\begin{aligned}
 A &= \operatorname{arctg} \frac{\zeta}{\xi}; \quad 0^\circ \leq A \leq 360^\circ; \\
 \gamma &= \operatorname{arsin} \frac{\eta}{D}; \quad -90^\circ \leq \gamma \leq 90^\circ; \\
 D &= \sqrt{(x - x_{\Pi})^2 + (y - y_{\Pi})^2 + (z - z_{\Pi})^2},
 \end{aligned} \tag{12}$$

где $x_{\Pi}, y_{\Pi}, z_{\Pi}$ — координаты пункта наблюдения, определяемые по формулам:

$$\begin{aligned}
 x_{\Pi} &= (r_B + h_{\Pi}) \cos B_{\Pi} \cos L_{\Pi}; \\
 y_{\Pi} &= (r_B + h_{\Pi}) \cos B_{\Pi} \sin L_{\Pi}; \\
 z_{\Pi} &= (r_B(1 - \alpha) + h_{\Pi}) \sin B_{\Pi}; \\
 r_B &= \frac{a_e}{\sqrt{1 - e^2 \cdot \sin^2(B)}},
 \end{aligned} \tag{13}$$

где $x_{\Pi}, y_{\Pi}, z_{\Pi}$ — координаты пункта наблюдения в ГСК; B_{Π}, L_{Π} — геодезические широта и долгота измерительного пункта соответственно; h_{Π} — высота пункта наблюдения над поверхностью общего земного эллипсоида; α — коэффициент полярного сжатия Земли; a_e — большая полуось Земного эллипсоида; e — эксцентриситет земного эллипсоида.

Входящие в формулы (12) величины ξ, η, ζ вычисляются по формулам:

$$\begin{aligned}
 \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} &= \Lambda \begin{bmatrix} x - x_{\Pi} \\ y - y_{\Pi} \\ z - z_{\Pi} \end{bmatrix}; \\
 \Lambda &= \begin{bmatrix} -\sin B_{\Pi} \cos L_{\Pi} & -\sin B_{\Pi} \sin L_{\Pi} & \cos B_{\Pi} \\ \cos B_{\Pi} \cos L_{\Pi} & \cos B_{\Pi} \sin L_{\Pi} & \sin L_{\Pi} \\ -\sin L_{\Pi} & \cos B_{\Pi} & 0 \end{bmatrix}.
 \end{aligned} \tag{14}$$

Согласно условию (5), определяющему скачки сопряженных переменных в моменты измерений, при решении навигационной задачи на основе вариационных условий оптимальности необходимо опре-

делять матрицу частных производных $\frac{\partial \bar{\psi}(\bar{q})}{\partial \bar{q}}$ от измеряемых параметров $\bar{\psi}$ по текущим кинематическим параметрам движения КО \bar{q} для вычисления скачков сопряженного вектора в каждой измерительной точке по следующему алгоритму:

$$\frac{\partial \bar{\psi}(\bar{q})}{\partial \bar{q}} = \left(\frac{\partial \bar{\psi}}{\partial \bar{r}}, \frac{\partial \bar{\psi}}{\partial \bar{V}} \right), \quad (15)$$

где

$$\frac{\partial \bar{\psi}}{\partial \bar{r}} = \frac{\partial \bar{\psi}}{\partial \bar{\rho}} \cdot \Lambda; \quad \frac{\partial \bar{\psi}}{\partial \bar{V}} = 0.$$

В этом выражении матрица $\frac{\partial \bar{\psi}}{\partial \bar{\rho}}$ определяется путем дифференцирования элементов вектора измеряемых параметров $\bar{\psi}$ по элементам вектора $\bar{\rho} = [\xi, \eta, \zeta]^T$ в топоцентрической ИСК.

Сопряженная система дифференциальных уравнений. Для решения навигационной задачи с использованием вариационных условий оптимальности необходимо конкретизировать вид системы дифференциальных уравнений, описывающих поведение сопряженного вектора $\bar{\lambda}$.

Если поведение исходной динамической системы описывается уравнениями общего вида (1), то сопряженные уравнения в системе (4) принимает следующий общий вид:

$$\dot{\bar{\lambda}} = - \left(\frac{\partial \bar{\Phi}}{\partial \bar{x}} \right)^T \bar{\lambda}. \quad (16)$$

В этом уравнении вектор $\bar{\Phi}$ соответствует вектору правых частей уравнений движения (8).

При записи сопряженной системы дифференциальных уравнений (16) необходимо учитывать возможность ее упрощения с точки зрения полноты учета малых возмущений, входящих в уравнения движения (8), на ограниченных мерных интервалах. К ним относятся сопротивление атмосферы и возмущающее ускорение, обусловленное влиянием сжатия Земли.

В известной работе В. С. Новоселова [26] было показано, что при решении задач оптимального управления на базе принципа максимума для уравнений движения объекта с малым параметром допустимо упрощать сопряженную систему уравнений без существенной потери точности с точки зрения достижения экстремального значения минимизируемого функционала. При этом уравнения движения КО сохраняют свой вид (8), в которых учитываются указанные возмущения.

Если при составлении сопряженной системы дифференциальных уравнений в уравнениях (8) пренебречь возмущающим влиянием атмосферы, то уравнения (16) можно представить следующим образом:

$$\begin{aligned}\dot{\lambda}_1 &= -(\Omega_3^2 - A)\lambda_4 - PE_\lambda x + 10BC \frac{xz}{r^2} \lambda_6; \\ \dot{\lambda}_2 &= -(\Omega_3^2 - A)\lambda_5 - PE_\lambda y + 10BC \frac{yz}{r^2} \lambda_6; \\ \dot{\lambda}_3 &= -(2BC - A)\lambda_6 - PE_\lambda z + 10BC \frac{z^2}{r^2} \lambda_6 + F_\lambda; \\ \dot{\lambda}_4 &= -\lambda_1 + 2\Omega_3 \lambda_5; \quad \dot{\lambda}_5 = -\lambda_2 - 2\Omega_3 \lambda_4; \quad \dot{\lambda}_6 = -\lambda_3,\end{aligned}\tag{17}$$

где

$$\begin{aligned}P &= 3A + 2BC(2D - 1); \\ Q &= 2BCD; \\ E_\lambda &= \frac{x\lambda_4 + y\lambda_5 + z\lambda_6}{r^2}; \\ F_\lambda &= QE_\lambda \frac{r^2}{z}.\end{aligned}$$

Особенностью этой системы сопряженных дифференциальных уравнений (17) состоит в том, что она строго учитывает сжатие ГПЗ. Данная система получена, согласно (16), путем дифференцирования правых частей уравнений движения (8) по элементам вектора фазового состояния КО x, y, z, V_x, V_y, V_z , а также простых алгебраических преобразований, связанных с приданием системе (17) наиболее компактной формы, удобной для программирования и экономичной в вычислительном отношении при проведении расчетов.

Если при составлении сопряженных систем дифференциальных уравнений дополнительно пренебречь влиянием сжатия ГПЗ, то урав-

нения (17) значительно упрощаются. В данном случае $\alpha_{20} = 0$ и выполняются следующие соотношения:

$$C = 0; \quad A = \alpha_{00}B; \quad P = 3A; \quad Q = 0.$$

Поэтому сопряженная система (17) принимает более простой вид:

$$\begin{aligned} \dot{\lambda}_1 &= -(\Omega_3^2 - B\alpha_{00})\lambda_4 - 3B\alpha_{00}E_\lambda x; \\ \dot{\lambda}_2 &= -(\Omega_3^2 - B\alpha_{00})\lambda_5 - 3B\alpha_{00}E_\lambda y; \\ \dot{\lambda}_3 &= B\alpha_{00}\lambda_5 - 3B\alpha_{00}E_\lambda z; \\ \dot{\lambda}_4 &= -\lambda_1 + 2\Omega_3\lambda_5; \quad \dot{\lambda}_5 = -\lambda_2 - 2\Omega_3\lambda_4; \quad \dot{\lambda}_6 = -\lambda_3. \end{aligned} \quad (18)$$

4. Краевая задача вариационного оценивания. На основании изложенного в пунктах 2 и 3 решения навигационной задачи в ГСК с использованием вариационных условий оптимальности сводится к решению двухточечной краевой задачи для П-системы из двенадцати нелинейных дифференциальных уравнений (8) и (17) либо (18). При использовании сопряженной системы (18) полная система дифференциальных уравнений, подлежащая совместному интегрированию в ходе решения краевой задачи, дается следующими соотношениями:

$$\begin{aligned} \dot{x} &= V_x; \quad \dot{y} = V_y; \quad \dot{z} = V_z; \\ \dot{V}_x &= (\Omega_3^2 - A)x + 2\Omega_3V_y - S\rho VV_x; \\ \dot{V}_y &= (\Omega_3^2 - A)y + 2\Omega_3V_x - S\rho VV_y; \\ \dot{V}_z &= (2BC - A)z - S\rho VV_z; \\ \dot{\lambda}_1 &= -(\Omega_3^2 - B\alpha_{00})\lambda_4 - 3B\alpha_{00}E_\lambda x; \\ \dot{\lambda}_2 &= -(\Omega_3^2 - B\alpha_{00})\lambda_5 - 3B\alpha_{00}E_\lambda y; \\ \dot{\lambda}_3 &= B\alpha_{00}\lambda_5 - 3B\alpha_{00}E_\lambda z; \\ \dot{\lambda}_4 &= -\lambda_1 + 2\Omega_3\lambda_5, \quad \dot{\lambda}_5 = -\lambda_2 - 2\Omega_3\lambda_4; \quad \dot{\lambda}_6 = -\lambda_3. \\ \lambda_s(t_0) &= 0; \quad \lambda_s(T) = 0; \quad s = 1, 2, \dots, 6. \end{aligned} \quad (19)$$

Здесь в уравнениях движения, описывающих поведение фазовых координат x, y, z, V_x, V_y, V_z , полностью учитываются как сжатие ГПЗ, так и влияние атмосферы.

В моменты измерений t_i сопряженный вектор $\bar{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_6]$ претерпевает скачкообразные изменения:

$$\bar{\lambda}(t_i^+) = \bar{\lambda}(t_i^-) + \Delta\bar{\lambda}(t_i), \quad (20)$$

где

$$\Delta\bar{\lambda} = \frac{\partial \bar{\Psi}_i^T}{\partial \bar{q}_i} K_{\delta_i}^{-1} \{ \bar{y}_i - \bar{\Psi}[\bar{q}(t_i)] \}.$$

Для рассматриваемого выше состава измерений значения $\bar{\Psi}[\bar{q}(t_i)]$ определяются по формулам (12), а элементы матрицы частных производных от измеряемых параметров по текущим элементам траектории — согласно (15).

В ходе решения краевой задачи необходимо определить неизвестный вектор исходного состояния \bar{q}_0 на заданный момент t_0 по известным краевым условиям для сопряженного вектора $\bar{\lambda}(t_0) = 0$ и $\bar{\lambda}(T) = 0$. Для решения этой задачи, как отмечено в пункте 2, могут применяться метод Ньютона (7), его модификации и другое.

5. Определение начального приближения для итерационного определения орбиты. Характеристики сходимости итерационных процессов вариационного оценивания параметров орбиты КО во многом зависят от точности определения начального приближения неизвестного вектора кинематических параметров. Начальные условия при итерационном решении задачи оценивания определяются в зависимости от того, имеются ли данные о конкретном КО в каталоге или он является новым. В первом случае для определения НУ необходимо проинтегрировать уравнение движения от момента времени, соответствующего данным каталога, до момента входа КО в зону видимости ОЭС или до начала момента времени принятого мерного интервала. Во втором случае для определения начального приближения могут быть, в частности, использованы приближенные алгоритмы, приведенные в работах [4, 24, 25], полученные применительно к условиям движения КО в центральном гравитационном поле Земли по трем или четырем измерительным точкам.

Различные алгоритмические способы решения данной задачи подробно рассмотрено в известной монографии М. Ф. Субботина [4]. Наиболее распространенным является метод Гаусса, который приводит к итерационному решению известного уравнения Лагранжа отно-

сительно дальности от измерительного пункта до КО. Для ускорения расчетов по методу Гаусса в этой работе, в частности, рекомендуется использовать приближенные формулы Гиббса.

Метод Гаусса широко применяется в астрономии для определения орбит астероидов и комет на гелиоцентрических орбитах. В этом случае сложно указать как дальности до наблюдаемых объектов, так и параметры их орбит. Однако, когда речь идет о системах контроля околоземного космического пространства, это не так. Практический интерес для таких систем представляют объекты искусственного происхождения, находящиеся на околокруговых и эллиптических геоцентрических орбитах.

Поэтому при разработке методов определения орбит таких КО необходимо учитывать имеющиеся реальные ограничения на топоцентрические дальности:

$$D_{\min} < D_i < D_{\max},$$

где $D_{\min} = 120$ км, $D_{\max} = 50000$ км и более.

Эти ограничения на топоцентрическую дальность определяются характеристиками орбит реально существующих КО. Их учет позволяет получить приближенные оценки дальности от измерительного пункта до КО. Методика использования метода Гаусса для определения начальных параметров движения КО по геоцентрическим орбитам изложена в [24]. Она предусматривает варьирование неизвестного значения начальной дальности в указанном диапазоне ее изменения при решении уравнений Лагранжа.

В настоящее время известно несколько модификаций метода Гаусса, которые используют одну и ту же процедуру для получения начального приближения и различаются в основном алгоритмами, используемыми в итерационных процедурах уточнения начального приближения. Такие модифицированные алгоритмы приведены в указанных выше работах [24, 25].

Из работ последних лет, связанных с данной проблематикой, отметим также работы [15-17]. В [15] и [16] рассмотрена задача начального определения орбиты КО по результатам оптических измерений, получаемых на совокупности коротких дуг, которые относят друг от друга на различные промежутки времени. В работе [17] предложена вероятностное определение начальной орбиты КО с использованием гауссовской смешанной модели.

Таким образом, в настоящее время имеется целый ряд алгоритмов приближенного определения начальных условий для оцени-

вания параметров орбитального движения КО по оптическим измерениям, которые могут обеспечить надежную сходимость вычислительного процесса решения рассматриваемой задачи в ее строгой нелинейной постановке.

6. Общий алгоритм решения краевой задачи вариационного оценивания. Общая последовательность вычислений при решении краевой задачи вариационного оценивания параметров орбитального движения КО состоит в следующем.

1. По предварительно определенному начальному значению неизвестного вектора фазового состояния КО $\bar{q}_0 = [\bar{r}_0, \bar{v}_0]^T$ и по начальным условиям для сопряженных переменных $\bar{\lambda}_0 = \bar{\lambda}(t_0) = 0$ совместно интегрируются основная и сопряженная системы дифференциальных уравнений П-системы (19) на принятом мерном интервале $t \in [t_0, T]$.

В ходе интегрирования в моменты измерений t_i , $i = 1(1)N$, по формулам (20) учитываются скачки сопряженных переменных $\Delta\lambda(t_i)$.

2. По результатам интегрирования дифференциальных уравнений, входящих в (19), вычисляются вектор состояния $\bar{q}(\bar{q}_0, T)$ и невязка для сопряженных переменных в конечный момент времени $\bar{\lambda}(\bar{q}_0, T)$.

3. Методом конечных разностей вычисляется значение и, затем, обращается матрица частных производных от сопряженных переменных в конечный момент времени T по определяемым начальным

условиям $\left[\frac{\partial \bar{\lambda}(\bar{q}_0, T)}{\partial \bar{q}_0} \right]_s$.

4. Реализуется вычислительная схема метода Ньютона по формуле (7) и определяется уточненное значение неизвестного вектора параметров движения \bar{q}_0 .

5. Проверяются заданные условия окончания итерационного процесса по точности решения. Если точностные требования не выполняются, то вычислительный процесс повторяется, начиная с пункта 1, при новых уточненных начальных условиях. Этот процесс повторяется итерационно, пока не будет достигнута требуемая точность расчета.

7. Результаты тестовых расчетов. Приведем некоторые результаты тестовых расчетов, связанных со статистическим оцениванием параметров орбиты по угловым измерениям ОЭС по приведенному в пункте 4 вариационному алгоритму метода максимального правдоподобия, применительно к КО МЕТЕОР – ПРИРОДА. Предполагалось, что ОЭС находится в районе горюда Кисловодск с геодезическими координатами $B_{II} = 43.91^\circ$, $L_{II} = 42.72^\circ$.

Основная цель численных расчетов заключалась в исследовании сходимости указанного выше алгоритма оценивания, а также влияния ошибок измерений на смещение получаемых оценок относительно «истинных» (точных) значений.

Исходные данные определялись по каталогу NORAD США в TLE-элементах [5] на дату 19.01.2017 года и время 03 часа 28 минут 59.82 секунд. Эти данные преобразовывались в прямоугольные элементы ГСК. Далее по полученным начальным условиям путем численного интегрирования уравнений движения решалась задача определения временных и кинематических параметров нахождения КО в зонах видимости заданной ОЭС. При этом были учтены требуемые условия освещенности КО (условия проведения измерений в ночное время), а также допустимые условия по углу возвышения КО над горизонтом ОЭС, который составлял 12° .

Для проведения тестовых расчетов была выбрана зона видимости КО МЕТЕОР – ПРИРОДА на дату 22.01.2017 года со временем входа в зону видимости 02 часа 47 минут 13 секунд и временем выхода из зоны 02 часа 53 минут 30 секунд. Таким образом, мерный интервал составил 377 секунд.

В качестве точных значений эфемерид КО принимались значения вектора состояния КО в ГСК, полученного на момент входа в указанную зону видимости.

На следующем этапе формировался массив угловых квазиизмерений. Для этого, во-первых, определялась расчетная траектория движения КО в зоне видимости на мерном интервале по параметрам которой производился расчет значений азимута и угла места на мерном интервале с дискретностью измерений 1 с. Для учета ошибок измерений проводилось статистическое моделирование выборки случайных измерений согласно нормальному закону их распределения с заданными значениями среднеквадратических отклонений $\sigma_A = 1''$ и $\sigma_\gamma = 1''$.

Сформированный массив случайных измерений на мерном интервале $T = 377$ с подвергался статистической обработке путем численного решения краевой задачи вариационного оценивания (19), (20) методом Ньютона (7) с использованием модели угловых измерений (12).

С целью проверки области сходимости решения выбирались достаточно грубые начальные приближения \bar{q}_{00} . Они отличались от точных значений \bar{q}_0 на 50 км по элементам вектора координат и на 50 м/с по проекциям вектора скорости. Итерационный процесс прекращался после выполнения пяти итераций.

Результаты расчетов приведены в таблицах 1 и 2.

Таблица 1. Результаты оптимального оценивания

Параметры оценивания	x_0 , км	y_0 , км	z_0 , км	V_{x0} , км/с	V_{y0} , км/с	V_{z0} , км/с
Точные значения	4643.812	3641.892	3561.745	-2.08926	-3.64065	6.43072
Начальное приближ-е	4593.812	3591.892	3511.745	-2.13926	-3.69065	6.38072
Оптимальн. оценки	4643.826	3641.896	3561.736	-2.08929	-3.64069	6.43077
Ошибки оценивания	0.013	0.005	-0.009	-0.00004	-0.00004	0.00005

Таблица 2. Сходимость вычислительного процесса

Номер итерации	Ошибки оценивания					
	δx_0 , м	δy_0 , м	δz_0 , м	δv_{x0} , м/с	δv_{y0} , м/с	δv_{z0} , м/с
0	50000	50000	50000	50	50	50
1	12451	4970	-9102	-031.79	-46.56	51.13
2	2085	997	-900	-4.42	-6.52	7.53
3	8	26	82	0.11	0.16	-0.11
4	30	-13	22	0.07	0.12	-0.13
5	10	5	-9	-0.04	-0.04	0.05

В таблице 1 даны точные значения параметров начального фазового состояния КА в абсолютной геоцентрической системе координат, принятое начальное приближение элементов уточняемого вектора, полученные в результате вариационной обработки измерений оптимальные оценки, а также характеристики точности оценивания.

В таблице 2 представлены значения абсолютных ошибок навигационного оценивания по итерациям. Анализ показывает, что в рассмотренных условиях достаточно проведения трех или четырех итераций. При выполнении четырех итераций абсолютные ошибки оценок вектора начального состояния КО не превышают 30 м по элементам вектора координат и 0.13 м/с по элементам вектора скорости. Максимальные относительные ошибки составляют по координатам не более $0.6 \cdot 10^{-5}$, а по элементам вектора скорости не более $0.3 \cdot 10^{-4}$.

Эти данные свидетельствуют о достаточно высокой точности и скорости сходимости вычислительного процесса даже при больших отклонениях начального приближения от точного значения оцениваемого вектора фазового состояния КО.

8. Заключение. В заключении отметим, что предлагаемые методические средства могут быть использованы при разработке и модернизации алгоритмов оптимального статистического оценивания

параметров состояния различных типов КО в составе автоматизированных комплексов обработки наблюдений и, в частности, при решении задачи определения параметров орбитального движения КО по угловым измерениям ОЭС, входящих в состав автоматизированной системы предупреждения об опасных ситуациях в околоземном космическом пространстве [12]. Они могут также применяться для решения задач тестирования приближенных алгоритмов навигационного оценивания и для выбора эффективной программы измерений.

В целом исследования показали, что вариационный подход является вполне работоспособным дополнением к существующей методологии навигационного определения параметров движения КО, основанной на прямом применении метода максимального правдоподобия. Предложенные алгоритмы могут применяться как самостоятельно, так и параллельно с алгоритмами прямого подхода для контроля правильности вычислений и обеспечения надежности результатов оценивания, что особенно важно в условиях массовых расчетов, проводимых в системе контроля космического пространства.

Литература

1. *Агаджанов П.А., Дулевич В.Е., Коростелева А.А.* Космические траекторные измерения // М.: Советское радио. 1969. 504 с.
2. *Нариманова Г.С. Тихонравов М.К.* Основы теории полета космических аппаратов // М.: Машиностроение. 1972. 608 с.
3. *Эльясберг П.Е.* Определение движения по результатам измерений // М.: Наука. 1976. 416 с.
4. *Субботин М.Ф.* Введение в теоретическую астрономию // М.: Наука. 1968. 800 с.
5. *Макарова Ю.Н.* Мониторинг техногенного засорения околоземного пространства и предупреждение об опасных ситуациях, создаваемых космическим мусором: Монография // М.: ЦНИИмаш. 2015. 244 с.
6. *Psiaki M.L.* Gaussian-Mixture Kalman Filter for Orbit Determination Using Angles-Only Data // Journal of Guidance, Control, and Dynamics. 2017. vol. 40. no. 9. pp. 2341–2347.
7. *DeMars K.J., Bishop R.H., Jah M.K.* Entropy-Based Approach for Uncertainty Propagation of Nonlinear Dynamical Systems // Journal of Guidance, Control, and Dynamics. 2013. vol. 36. no. 4. pp. 1047–1057.
8. *DeMars K.J., Jah M.K.* Probabilistic Initial Orbit Determination Using Gaussian Mixture Models // Journal of Guidance, Control, and Dynamics. 2013. vol. 36. no. 5. pp. 1324–1335.
9. *Horwood J.T., Poore A.B.* Gauss von Mises Distribution for Improved Uncertainty Realism in Space Situational Awareness // SIAM/ASA Journal of Uncertainty Quantification. 2014. vol. 2. no. 1. pp. 276–304.
10. *Jones B.A., Doostan A., Born G.H.* Nonlinear Propagation of Orbit Uncertainty Using Non-Intrusive Polynomial Chaos // Journal of Guidance, Control, and Dynamics. 2013. vol. 36. no. 2. pp. 430–444.
11. *Bayramoglu E., Ravn O.* A Novel Hypothesis Splitting Method Implementation for Multi-Hypothesis Filters // 2013 10th IEEE International Conference on Control and Automation (ICCA). 2013. pp. 574–579.

12. *Psiaki M.* Gaussian Mixture Nonlinear Filtering with Resampling for Mixand Narrowing // IEEE Transactions on Signal Processing. 2016. vol. 64. no. 21. pp. 5499–5512.
13. *Psiaki M.L., Schoenberg J.R., Miller I.T.* Gaussian Sum Reapproximation for Use in a Nonlinear Filter // Journal of Guidance, Control, and Dynamics. 2015. vol. 38. no. 2. pp. 292–303.
14. *Psiaki M.L., Weisman R.M., Jah M.K.* Gaussian Mixture Approximation of Angles-Only Initial Orbit Determination Likelihood Function // Journal of Guidance, Control, and Dynamics. 2017. vol. 40. no. 11. pp. 2807–2819.
15. *Lei X. et al.* A geometrical approach to association of space-based very short-arc LEO tracks // Advances in Space Research. 2018. vol. 62. no. 3. pp. 542–553.
16. *Cail H. et al.* Improved tracklet association for space objects using short-arc optical measurements // Acta Astronautica. 2018. vol. 151. pp. 836–847.
17. *DeMars K.J., Jah M.* Probabilistic Initial Orbit Determination Using Gaussian Mixture Models // Journal of Guidance, Control, and Dynamics. 2013. vol. 36. no. 5. pp. 1324–1335.
18. *Шугаров А.С. и др.* О концепции экономичной космической системы обнаружения опасных небесных тел // Космические исследования. 2015. Т. 53. № 2. С. 95–104.
19. *Ахметшин П.З.* Космический патруль: варианты схемы оптического барьера // Космические исследования. 2013. Т. 51. № 4. С. 274–286.
20. *Нароенков С.А., Шустов Б.М., Емельяненко В.В.* О длине дуги наблюдения малого тела солнечной системы, достаточной для классификации его как опасно-го // Космические исследования. 2013. Т. 51. № 5. С. 372–379.
21. *Марков Ю.Г. и др.* Анализ влияния различных возмущающих факторов на высокоточный прогноз орбит космических аппаратов // Космические исследования. 2016. Т. 54. № 2. С. 164–172.
22. *Миронов В.И., Миронов Ю.В.* Комплексное оценивание параметров состояния нелинейных динамических систем на основе вариационного подхода // Труды СПИИРАН. 2008. Вып. 7. С. 222–229.
23. *Mironov V., Mironov Y., Sokolov B., Yusupov R.* Variational maximum-likelihood method and its application in the aerospace navigation // Information technology and management science. 2013. vol. 16. no. 1. pp. 27–32.
24. *Самтохин А.С., Хуторовский З.Н.* Метод первоначального определения параметров околоземных орбит по трем угловым измерениям // Препринты Института прикладной математики им. М.В. Келдыша РАН. 2014. № 0. С. 44–31.
25. *Шеффер В.А.* Вычисление промежуточной возмущенной орбиты по трем и более положениям малого тела на небесной сфере // Астрономический вестник. 2013. Т. 47. С. 40–52.
26. *Новоселов В.С.* Аналитическая теория оптимизации в гравитационных полях // ЛГУ. 1972. 317 с.

Миронов Вячеслав Иванович — д-р техн. наук, профессор, ведущий научный сотрудник, лаборатория информационных технологий в системном анализе и моделировании, Федеральное государственное бюджетное учреждение науки Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: фундаментальные и прикладные исследования проблем комплексного моделирования, теория оптимального наблюдения и управления динамическими процессами, вычислительная математика, баллистика космических полетов, методы навигации и управления движением ракет-носителей и космических аппаратов, статистический анализ характеристик сложных технических систем. Число научных публикаций — 350. mironuv@yandex.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Российская Федерация; р.т.: +7(812)328-3337; факс: +7(812)328-4450.

Миронов Юрий Вячеславович — д-р техн. наук, доцент, ведущий специалист, отдел системного проектирования, АО "Научно-инженерный центр Санкт-Петербургского электро-технического университета"(АО "НИЦ СПб ЭТУ"). Область научных интересов: баллистика космических полетов, методы навигации и управления движением ракет-носителей и космических аппаратов, теория статистического оценивания, вычислительная математика. Число научных публикаций — 150. mironoww.yuriy@yandex.ru; ул. Политехническая, 22, лит. Н, 194021, Санкт Петербург, Российская Федерация; р.т.: +7(812) 703-7583; факс: +7(812) 703-7584.

Хегай Дмитрий Климович — канд. техн. наук, доцент, директор департамента, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики (Университета «ИТМО»). Область научных интересов: ракетно-космические технологии, космические инфотелекоммуникационные технологии, оптоэлектронные технологии, экспериментальные исследования в космосе. Число научных публикаций — 70. hdk_ifmo@mail.ru; ул. Политехническая, 49, 197101, Санкт Петербург, Российская Федерация; р.т.: +7(812) 437-1831; факс: +7(812) 498-1070.

V.I. MIRONOV, Y.V. MIRONOV, D.K. KHEGAY
**OPTIMAL DETERMINATION OF SPACE OBJECTS ORBIT BY
ANGULAR MEASUREMENTS OF GROUND-BASED
OPTOELECTRONIC STATIONS**

Mironov V.I., Mironov Y.V., Khegay D.K. **Optimal Determination of Space Objects Orbit by Angular Measurements of Ground-Based Optoelectronic Stations**

Abstract. The successful solution of practical cosmonautics problems is largely achieved by contemporary advances in measurement and computing technology, as well as by improvements in methods of primary and secondary processing of trajectory measurements. Therefore, in long-range programs of space exploration and space technology development, much attention is paid to improving existing and developing new algorithmic and technical means of navigation support for flights of space objects with the purpose to expand capabilities and increasing the efficiency of autonomous navigation systems of spacecraft, as well as ground-based and perspective orbital systems of space monitoring. Currently, active work is underway to modernize and develop promising complexes of specialized optoelectronic devices for monitoring near-Earth space based on angular measurements. The article considers the application of the variational approach for solving problems of statistical estimation of the trajectory parameters of the orbital object by angular measurements, which were carried out by ground-based optoelectronic means that are part of the modern space control system. Models and algorithms for determining estimates of orbital parameters that implement the variational version of the maximum likelihood method are presented, as well as the results of test calculations related to iterative solution of the two-point boundary value problem of variational estimation. The main purpose of the numerical calculations is a study of convergence of the proposed estimation algorithm, as well as the impact of measurement errors on the displacement of the obtained estimates relative to their exact values. The simulation results, presented in the article, correspond to the conditions of the orbital motion of METEOR PRIRODA spacecraft and were obtained using the ephemeris data of the NORAD catalog in TLE-elements.

Keywords: Statistical Estimation, Spacecraft, Nonlinear Dynamical Systems, Maximum Likelihood Criterion, Optical Angular Measurements.

Mironov Vyacheslav Ivanovich — Ph.D., Dr.Sci., Professor, Leading Researcher, Laboratory of Information Technologies in System Analysis and Modeling, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS). Research interests: fundamental and applied researches in complex modeling; the theory of optimal observation and control of dynamic processes; computing mathematics; the ballistics of space flight; statistical analysis of complex technical systems. The number of publications — 350. mironuv@yandex.ru; 39, 14-th Line V.O., 199178, St. Petersburg, Russian Federation; office phone: +7(812)328-3337; fax: +7(812)328-4450.

Mironov Yuri Vyacheslavovich — Ph.D., Dr.Sci., Associate Professor, Leading Researcher, Department of System Design, JSC "R&EC ETU". Research interests: ballistics of space flights, methods of navigation and control of the motion of carrier rockets and space vehicles, the theory of statistical estimation, computational mathematics. The number of publications — 150. mironoww.yuriy@yandex.ru; 22, лит. Н, Politechnicheskaya str., 194021, St. Petersburg, Russian Federation; office phone: +7(812) 703-7583; fax: +7(812) 703-7584.

Khegay Dmitry Klimovich — Ph.D., Associate Professor, Department Director, ITMO University (Saint Petersburg National Research University of Information Technologies, Mechan-

ics and Optics). Research interests: space-rocket technology, space info-telecommunications technologies, optoelectronic technology, experimental research in space. The number of publications — 70. hdk_ifmo@mail.ru; 49, Kronverksky pr., 197101, St. Petersburg, Russian Federation; office phone: +7(812) 437-1831; fax: +7(812) 498-1070.

References

1. Agadzhanyan P.A., Dulevich V.E., Korostelev A.A. *Kosmicheskie traektornye izmerenija* [Space trajectory measurements]. M.: Sov. radio. 1969. 504 p. (In Russ.).
2. Narimanov G.S., Tihonravova M.K. *Osnovy teorii poleta kosmicheskikh apparatov* [The fundamentals of the theory of flight of space vehicles]. M.: Mashinostroenie. 1972. 608 p. (In Russ.).
3. Jel'jasberg P.E. *Opreделение dvizhenija po rezul'tatam izmerenij* [Determination of motion from measurements]. M.: Nauka. 1976. 416 p.
4. Subbotin M.F. *Vvedenie v teoreticheskuyu astronomiju* [Introduction to theoretical astronomy]. M.: Nauka. 1968. 800 p. (In Russ.).
5. Makarova Ju.N. *Monitoring tehnogennogo zasorenija okolozemnogo prostranstva i preduprezhdenie ob opasnykh situacijah, sozdavaemyh kosmicheskim musorom* [Monitoring technogenic pollution of near-Earth space and warning about dangerous situations created by space debris]. M.: CNIImash. 2015. 244 p. (In Russ.).
6. Psiaki M.L. Gaussian-Mixture Kalman Filter for Orbit Determination Using Angles-Only Data. *Journal of Guidance, Control, and Dynamics*. 2017. vol. 40. no. 9. pp. 2341–2347.
7. DeMars K.J., Bishop R.H., Jah M.K. Entropy-Based Approach for Uncertainty Propagation of Nonlinear Dynamical Systems. *Journal of Guidance, Control, and Dynamics*. 2013. vol. 36. no. 4. pp. 1047–1057.
8. DeMars K.J., Jah M.K. Probabilistic Initial Orbit Determination Using Gaussian Mixture Models. *Journal of Guidance, Control, and Dynamics*. 2013. vol. 36. no. 5. pp. 1324–1335.
9. Horwood J.T., Poore A.B. Gauss von Mises Distribution for Improved Uncertainty Realism in Space Situational Awareness. *SIAM/ASA Journal of Uncertainty Quantification*. 2014. vol. 2. no. 1. pp. 276–304.
10. Jones B.A., Doostan A., Born G.H. Nonlinear Propagation of Orbit Uncertainty Using Non-Intrusive Polynomial Chaos. *Journal of Guidance, Control, and Dynamics*. 2013. vol. 36. no. 2. pp. 430–444.
11. Bayramoglu E., Ravn O. A Novel Hypothesis Splitting Method Implementation for Multi-Hypothesis Filters. 2013 10th IEEE International Conference on Control and Automation (ICCA). 2013. pp. 574–579.
12. Psiaki M. Gaussian Mixture Nonlinear Filtering with Resampling for Mixand Narrowing. *IEEE Transactions on Signal Processing*. 2016. vol. 64. no. 21. pp. 5499–5512.
13. Psiaki M.L., Schoenberg J.R., Miller I.T. Gaussian Sum Reapproximation for Use in a Nonlinear Filter. *Journal of Guidance, Control, and Dynamics*. 2015. vol. 38. no. 2. pp. 292–303.
14. Psiaki M.L., Weisman R.M., Jah M.K. Gaussian Mixture Approximation of Angles-Only Initial Orbit Determination Likelihood Function. *Journal of Guidance, Control, and Dynamics*. 2017. vol. 40. no. 11. pp. 2807–2819.
15. Lei X. et al. A geometrical approach to association of space-based very short-arc LEO tracks. *Advances in Space Research*. 2018. vol. 62. no. 3. pp. 542–553.
16. Cail H. et al. Improved tracklet association for space objects using short-arc optical measurements. *Acta Astronautica*. 2018. vol. 151. pp. 836–847.
17. DeMars K.J., Jah M. Probabilistic Initial Orbit Determination Using Gaussian Mixture Models. *Journal of Guidance, Control, and Dynamics*. 2013. vol. 36. no. 5. pp. 1324–1335.

18. Shugarov A.S. et al. [On the concept of an economical space system for detecting dangerous celestial bodies]. *Kosmicheskie Issledovaniya – Cosmic Research*. 2015. Issue 53. vol. 2. pp. 95–104. (In Russ.).
19. Ahmetshin R.Z. [Space Patrol: Optical Barrier Options]. *Kosmicheskie Issledovaniya – Cosmic Research*. 2013. Issue 51. vol. 4. pp. 274–286. (In Russ.).
20. Naroenkov S.A., Shustov B.M., Emel'yanenko V.V. [On the length of the arc of observation of a small body of the solar system, sufficient to classify it as dangerous]. *Kosmicheskie Issledovaniya – Cosmic Research*. 2013. Issue 51. vol. 5. pp. 372–379. (In Russ.).
21. Markov Ju.G. et al. [Analysis of the influence of various disturbing factors on the high-precision prediction of spacecraft orbits]. *Kosmicheskie Issledovaniya – Cosmic Research*. 2016. Issue 54. vol. 2. pp. 164–172. (In Russ.).
22. Mironov V.I., Mironov Ju.V. [Complex estimation of the state parameters of nonlinear dynamical systems on the basis of the variational approach]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2008. vol. 7. pp. 222–229. (In Russ.).
23. Mironov V., Mironov Y., Sokolov B., Yusupov R. Variational maximum-likelihood method and its application in the aerospace navigation. *Information technology and management science*. 2013. vol. 16. no. 1. pp. 27–32.
24. Samotohin A.S., Hutorovskij Z.N. [Method of initial determination of parameters of near-earth orbits with three angular measurements]. *Preprinty Instituta prikladnoj matematiki im. M.V.Keldysha RAN Keldysh Institute of Applied Mathematics*. 2014. vol. 0. pp. 44–31. (In Russ.).
25. Sheffer. V.A. [Calculation of the intermediate perturbed orbit in three or more positions of a small body on the celestial sphere]. *Astronomicheskij vestnik – Solar System Research*. 2013. vol. 47. pp. 40–52. (In Russ.).
26. Novoselov, V.S. *Analiticheskaya teoriya optimizacii v gravitacionnyh polyah* [Analytical theory of optimization in gravitational fields]. LGU. 1972. 317 p. (In Russ.).

Signed to print 25.09.2019

Printed in Publishing center GUAP, 67, B. Morskaya, St. Petersburg, 190000, Russia

The journal is registered in Russian Federal Agency for Communications
and Mass-Media Supervision, certificate ПИ № ФС77-41695 dated August 19, 2010 г.
Subscription Index П5513, Russian Post Catalog

Подписано к печати 25.09.2019. Формат 60×90 1/16. Усл. печ. л. 12,84. Заказ № 398.

Тираж 150 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП, 190000, Санкт-Петербург, Б. Морская, д. 67

Журнал зарегистрирован Федеральной службой по надзору в сфере связи
и массовых коммуникаций,
свидетельство ПИ № ФС77-41695 от 19 августа 2010 г.

Подписной индекс П5513 по каталогу «Почта России»

РУКОВОДСТВО ДЛЯ АВТОРОВ

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Труды СПИИРАН» <http://www.proceedings.spiiras.nw.ru>. При регистрации авторам рекомендуется заполнить все предложенные поля данных.

Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше. Объем основного текста – от 20 до 30 страниц включительно. Формат страницы документа – А5 (148 мм ширина, 210 мм высота); ориентация – портретная; все поля – 20 мм. Верхний и нижний колонтитулы страницы – пустые. Основной шрифт документа – Times New Roman, основной кегль (размер) шрифта – 10 pt. Переносы разрешены. Абзацный отступ устанавливается размером в 10 мм. Межстрочный интервал – одинарный. Номера страниц не проставляются.

В основную часть допускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

AUTHOR GUIDELINES

Interaction between each potential author and the Editorial board is realized through the personal account on the website of the journal "SPIIRAS Proceedings" <http://www.proceedings.spiiras.nw.ru>. At the registration the authors are requested to fill out all data fields in the proposed form.

The submissions should be prepared using MS Word 2007 text editor or higher versions, at that, only manuscripts in *.docx format will be considered. The text of the paper in the main part of it should be from 20 – 30 pages of A5 size that is 210 X 148 mm; orientation – portrait; all margins – 20 mm. The font of the main paper text is Times New Roman of 10 pt font size. The pages' headers and footers should be empty; indentation – 10 mm; line spacing – single; pages are not numbered; hyphenations are allowed.

Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered by the paper template in more detail in journal web.

ISSN 2078-9181



9 772078 918785 >

