

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

5(78)/2015

5(78)/2015

REFEREED EDITION

INFORMATSIONNO- UPRAVLIAYUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder

«Information and Control Systems», Ltd.

PublisherSaint-Petersburg State University
of Aerospace Instrumentation**Editor-in-Chief**

M. Sergeev

Dr. Sc., Tech., Professor, St. Petersburg, Russia

Deputy Editor-in-Chief

E. Krouk

Dr. Sc., Tech., Professor, St. Petersburg, Russia

Executive secretary

O. Muravtsova

Editorial Council

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

L. Chubraeva

RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Tech., Professor, St. Petersburg, Russia

V. Kozlov

Dr. Sc., Tech., Professor, St. Petersburg, Russia

B. Meyer

Dr. Sc., Professor, Zurich, Switzerland

A. Ovodenko

Dr. Sc., Tech., Professor, St. Petersburg, Russia

Y. Podoplyokin

Dr. Sc., Tech., Professor, St. Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Phys.-Math., Novosibirsk, Russia

V. Simakov

Dr. Sc., Tech., Professor, Moscow, Russia

V. Vasilev

RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia

R. Yusupov

RAS Corr. Member, Dr. Sc., Tech., Professor, St. Petersburg, Russia

Editorial Board

V. Anisimov

Dr. Sc., Tech., Professor, St. Petersburg, Russia

B. Bezruchko

Dr. Sc., Phys.-Math., Saratov, Russia

N. Blaunstein

Dr. Sc., Phys.-Math., Professor, Beer-Sheva, Israel

A. Dudin

Dr. Sc., Tech., Professor, Minsk, Belarus

I. Dumer

PhD., Professor, Riverside, USA

V. Khimenko

Dr. Sc., Tech., Professor, St. Petersburg, Russia

G. Maltsev

Dr. Sc., Tech., Professor, St. Petersburg, Russia

V. Melekhin

Dr. Sc., Tech., Professor, St. Petersburg, Russia

A. Shalyto

Dr. Sc., Tech., Professor, St. Petersburg, Russia

A. Shepeta

Dr. Sc., Tech., Professor, St. Petersburg, Russia

A. Smirnov

Dr. Sc., Tech., Professor, St. Petersburg, Russia

Z. Yuldashev

Dr. Sc., Tech., Professor, St. Petersburg, Russia

A. Zeifman

Dr. Sc., Phys.-Math., Vologda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** N. Karavaeva, M. Chernenko**Layout and composition:** N. Karavaeva**Contact information**

The Editorial and Publishing Center, SUAI

67, B. Morskaia, 190000, St. Petersburg, Russia

Website: <http://i-us.ru/en>, E-mail: ius.spb@gmail.com

Tel.: +7 - 812 494 70 02

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

THEORETICAL AND APPLIED MATHEMATICS**Balonin N. A., Djokovic D. Z.** Negaperiodic Golay Pairs
and Hadamard Matrices

2

INFORMATION PROCESSING AND CONTROL**Krasilnikov N. N., Krasilnikova O. I.** Methods of Converting 2D Images
and Videos into Stereoscopic Format

18

Lyamin A. V., Cherepovskaya E. N. Experimental Studies of Biometric
User Identification Based on the Data of Eye Tracker Tobii X2-30

26

Yastrebov V. A., Veselov A. I., Gilmudinov M. R. Basic Methods and
Principles of Developing Image Inpainting Algorithms

34

INFORMATION AND CONTROL SYSTEMS**Khimenko V. I.** Laser Information Systems: Developing a Statistical Theory

43

Ardashov A. A., Arseniev V. N., Silantsev S. B. Processing Miscellaneous
Statistical Information about Control System Accuracy Characteristics

55

Muromtsev D. Yu., Gribkov A. N., Kurkin I. A. Knowledge Base of an
Intelligent Information Control System for a Multidimensional Object

60

SYSTEM AND PROCESS MODELING**Barkovsky E. A., Sokolov A. V.** Optimal Control of Two Parallel FIFO
Queues on an Infinite Time

65

Grankin M. A. Model for Estimating the Average Number of Backlogged
Users of M2M Systems in 5G Mobile Networks

72

HARDWARE AND SOFTWARE RESOURCES**Kosmynina N. A., Legalov A. I., Barkov A. V., Lapin A. A.** Spacecraft
Control Language Development

82

Ponomarev A. V. Decentralized Recommendation System Architecture
Based on Locality-Sensitive Hashing

91

INFORMATION CODING AND TRANSMISSION**Malichenko D. A.** Heuristic Algorithm for Calculating the Size of Tiers
in a Hierarchical Storage System

100

INFORMATION CHANNELS AND MEDIUM**Gurevich V. E., Egorov S. G.** Performance Analysis of Nonlinear
Distortions of RF Signal in Wireless CDMA Downlink with Quadrature
Amplitude Modulation

106

SYSTEM ANALYSIS**Vostrikov A. A., Sergeev M. B., Litvinov M. Yu.** Masking of Digital
Visual Data: the Term and Basic Definitions

116

CONTROL IN SOCIAL AND ECONOMIC SYSTEMS**Grigorieva X. V.** Pareto Optimality in a Static Competitive
Decision-Making Model

124

INFORMATION ABOUT THE AUTHORS

130

Submitted for publication 03.09.15. Passed for printing 20.10.15. Format 60×84_{1/8}.
Offset paper. Phototype SchoolBookC. Offset printing.

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

The journal is distributed by subscription. Subscription can be made in the Editorial and publishing
center, SUAI as well as in any post office based on «Rospechat» catalogue:
№ 48060 — annual subscript, № 15385 — semiannual subscript.

5(78)/2015

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Учредитель
ООО «Информационно-управляющие системы»Издатель
Санкт-Петербургский государственный университет
аэрокосмического приборостроенияГлавный редактор
М. Б. Сергеев,
д-р техн. наук, проф., С.-Петербург, РФЗам. главного редактора
Е. А. Крук,
д-р техн. наук, проф., С.-Петербург, РФОтветственный секретарь
О. В. МуравцоваРедакционный совет:
Председатель А. А. Оводенко,
д-р техн. наук, проф., С.-Петербург, РФ
В. Н. Васильев,
чл.-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФ
В. Н. Козлов,
д-р техн. наук, проф., С.-Петербург, РФ
К. Кристоделу,
д-р наук, проф., Альбукерке, Нью-Мексико, США
Б. Мейер,
д-р наук, проф., Цюрих, Швейцария
Ю. Ф. Подоплёкин,
д-р техн. наук, проф., С.-Петербург, РФ
В. В. Симаков,
д-р техн. наук, проф., Москва, РФ
Л. Фортун,
д-р наук, проф., Катания, Италия
А. Л. Фрадков,
д-р техн. наук, проф., С.-Петербург, РФ
Л. И. Чубраева,
чл.-корр. РАН, д-р техн. наук, С.-Петербург, РФ
Ю. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФ
Р. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., С.-Петербург, РФРедакционная коллегия:
В. Г. Анисимов,
д-р техн. наук, проф., С.-Петербург, РФ
Б. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФ
Н. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, Израиль
А. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, Беларусь
И. И. Думер,
д-р наук, проф., Риверсайд, США
А. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФ
Г. Н. Мальцев,
д-р техн. наук, проф., С.-Петербург, РФ
В. Ф. Мелехин,
д-р техн. наук, проф., С.-Петербург, РФ
А. В. Смирнов,
д-р техн. наук, проф., С.-Петербург, РФ
В. И. Хименко,
д-р техн. наук, проф., С.-Петербург, РФ
А. А. Шалыто,
д-р техн. наук, проф., С.-Петербург, РФ
А. П. Шепета,
д-р техн. наук, проф., С.-Петербург, РФ
З. М. Юлдашев,
д-р техн. наук, проф., С.-Петербург, РФРедактор: А. Г. Ларионова
Корректор: Т. В. Звертановская
Дизайн: Н. Н. Караваева, М. Л. Черненко
Компьютерная верстка: Н. Н. КараваеваАдрес редакции: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Тел.: (812) 494-70-02, e-mail: ius.spb@gmail.com, сайт: <http://i-us.ru>Журнал зарегистрирован в Министерстве РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-49181 от 30 марта 2012 г.Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий,
в которых должны быть опубликованы основные научные результаты диссертации
на соискание ученой степени доктора и кандидата наук».

© Коллектив авторов, 2015

ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ МАТЕМАТИКА

Balonin N. A., Djokovic D. Z. Negaperiodic Golay Pairs and Hadamard Matrices 2

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

Красильников Н. Н., Красильникова О. И. Методы конвертации 2D-изображений и видео в стереоскопический формат 18*Лямин А. В., Череповская Е. Н. Экспериментальные исследования биометрической идентификации пользователей на основе данных айтрекера Tobii X2-30* 26*Ястребов В. А., Веселов А. И., Гильмутдинов М. Р. Базовые методы и принципы построения алгоритмов восстановления регионов изображений* 34

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Хименко В. И. Лазерные информационные системы: принципы построения статистической теории 43*Ардашов А. А., Арсеньев В. Н., Силантьев С. Б. Метод обработки неоднородной статистической информации о характеристиках точности системы управления* 55*Муромцев Д. Ю., Грибков А. Н., Куркин И. А. База знаний интеллектуальной информационно-управляющей системы многомерным объектом* 60

МОДЕЛИРОВАНИЕ СИСТЕМ И ПРОЦЕССОВ

Барковский Е. А., Соколов А. В. Оптимальное управление двумя параллельными FIFO-очередями на бесконечном времени 65*Гранкин М. А. Модель для оценки среднего количества активных абонентов в системе межмашинного взаимодействия в сетях 5-го поколения* 72

ПРОГРАММНЫЕ И АППАРАТНЫЕ СРЕДСТВА

Космынина Н. А., Легалов А. И., Барков А. В., Лапин А. А. Разработка языка управления космическими аппаратами 82*Пономарев А. В. Архитектура децентрализованной рекомендующей системы, основанной на применении локально-чувствительного хеширования* 91

КОДИРОВАНИЕ И ПЕРЕДАЧА ИНФОРМАЦИИ

Маличенко Д. А. Эвристический алгоритм расчета размеров памяти в многоуровневой системе хранения 100

ИНФОРМАЦИОННЫЕ КАНАЛЫ И СРЕДЫ

Гуревич В. Э., Егоров С. Г. Исследование нелинейных искажений радиосигнала в системе CDMA с квадратурно-амплитудной модуляцией 106

СИСТЕМНЫЙ АНАЛИЗ

Востриков А. А., Сергеев М. Б., Литвинов М. Ю. Маскирование цифровой визуальной информации: термин и основные определения 116

УПРАВЛЕНИЕ В СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМАХ

Григорьева К. В. Парето-оптимальность в статической конкурентной модели принятия решений 124

СВЕДЕНИЯ ОБ АВТОРАХ

130

Сдано в набор 03.09.15. Подписано в печать 20.10.15. Формат 60×84 1/8.
Бумага офсетная. Гарнитура SchoolBookC. Печать офсетная.
Усл. печ. л. 15,6. Уч.-изд. л. 19,5. Тираж 1000 экз. Заказ 377.Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.
190000, Санкт-Петербург, Б. Морская ул., 67.Журнал распространяется по подписке. Подписку можно оформить
через редакцию, а также в любом отделении связи по каталогу «Роспечать»:
№ 48060 — годовой индекс, № 15385 — полугодовой индекс.

NEGAPERIODIC GOLAY PAIRS AND HADAMARD MATRICES

N. A. Balonin^a, Dr. Sc., Tech., Professor, korbendfs@mail.ru

D. Z. Djokovic^b, PhD, Distinguished Professor Emeritus, djokovic@uwaterloo.ca

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

^bUniversity of Waterloo, Department of Pure Mathematics and Institute for Quantum Computing, Waterloo, Ontario, N2L 3G1, Canada

Purpose: In analogy with the ordinary and the periodic Golay pairs, we introduce also the negaperiodic Golay pairs. (They occurred first, under a different name, in a paper of Ito.) **Methods:** We investigate the construction of Hadamard (and weighing) matrices from two negacyclic blocks (2N-type). The Hadamard matrices of 2N-type are equivalent to negaperiodic Golay pairs. **Results:** If a Hadamard matrix is also a Toeplitz matrix, we show that it must be either cyclic or negacyclic. We show that the Turyn multiplication of Golay pairs extends to a more general multiplication: one can multiply Golay pairs of length g and negaperiodic Golay pairs of length v to obtain negaperiodic Golay pairs of length gv . We show that the Ito's conjecture about Hadamard matrices is equivalent to the conjecture that negaperiodic Golay pairs exist for all even lengths. **Practical relevance:** Hadamard matrices have direct practical applications to the problems of noise-immune coding and compression and masking of video information.

Keywords — Hadamard Matrices, Cyclic Matrices, Negacyclic Matrices, Periodic Golay Pairs, Negaperiodic Golay Pairs.

Introduction

The Golay pairs (abbreviated as G-pairs, and also known as Golay sequences) have been introduced in a note of M. Golay [1] published in 1961. Since then they have been studied by many researchers and used in various combinatorial constructions, in particular for the construction of Hadamard matrices [2] and [3, Ch. 23].

The periodic Golay pairs (PG-pairs) made their first appearance, under a different name, in a note of the second author [4] published in 1998. They are equivalent to Hadamard matrices built from two circulant blocks (2C-type). It is now known that periodic Golay pairs exist for infinitely many lengths for which no ordinary Golay pairs are known [5].

In this paper we complete the picture by defining the negaperiodic Golay pairs (NG-pairs). These pairs are equivalent to Hadamard matrices built from two negacyclic blocks (2N-type). The NG-pairs were first introduced by N. Ito, under the name of “associated pairs”, in his paper [6] published in 2000. An interesting observation is that the ordinary Golay pairs are precisely the pairs which are both PG and NG-pairs.

In an earlier paper [7] Ito proposed a conjecture which is stronger than the famous Hadamard conjecture. It turns out that his conjecture is equivalent to the assertion that the NG-pairs exist for all even lengths. This is drastically different from the known facts about ordinary and periodic Golay pairs. Examples of NG-pairs of even length ≤ 92 are listed in [6]. As far as we know, no NG-pairs of length 94 have been constructed.

We now describe the content of each of the remaining sections.

k- Toeplitz matrices: We show that if a Hadamard matrix is also a Toeplitz matrix, then it must be cyclic or negacyclic. As cyclic Hadamard matrices beyond order 4 are not likely to exist, we conjecture that the same holds true for negacyclic Hadamard matrices beyond order 2. We have verified the latter conjecture for orders ≤ 40 . As a substitute for Ito's conjecture we propose the weaker conjecture in which the two negacyclic blocks are replaced by Toeplitz matrices.

Three kind of Golay pairs: We define negaperiodic autocorrelation function (NAF) and NG-pairs. These are binary sequences of the same length v whose NAFs add up to zero. The length v must be an even integer or 1. For the sake of comparison we recall some facts about ordinary and PG-pairs. We show that the Turyn multiplication of G-pairs extends to give a multiplication of G-pairs and NG-pairs. More precisely, one can multiply G-pairs of length g and NG-pairs of length v to obtain NG-pairs of length gv . In particular, one can double the length of any NG-pair. We also define a natural equivalence relation for NG-pairs.

Cyclic relative difference families: We introduce a natural bijection Φ_v from the set of binary sequences of length v onto the set of v -subsets of $\mathbb{Z}_{2v}$. We recall the definition of the relative difference families in the cyclic group $\mathbb{Z}_{2v}$ with respect to the subgroup of order 2. We show that a pair of binary sequences of length v is an NG-pair if and only if the Φ_v -images of these sequences form a relative difference family in $\mathbb{Z}_{2v}$. We also show that Ito's conjecture, which entails the Hadamard matrix conjecture, is equivalent to the assertion that NG-pairs exist for all even lengths v .

There are only a few known infinite series of NG-pairs. In the subsequent three sections we treat two of them, the first and second Paley series. First we recall the definition of Paley conference matrices (C-matrices). They have order $1 + q$ where q is an odd prime power. Those for $q \equiv 1 \pmod{4}$ give rise to the first Paley series of NG-pairs, with length $1 + q$. Those for $q \equiv 3 \pmod{4}$ give rise to the second Paley series of NG-pairs, with length $(1 + q)/2$. The main facts that we use are that all Paley C-matrices of the same order are equivalent and that each of these equivalence classes contains a negacyclic C-matrix.

Ito series: We recall that Ito constructed in [7] an infinite series of relative difference sets in dicyclic groups. Hence, this gives an infinite series of NG-pairs to which we refer as the Ito series. However, we show that the Ito series is contained in the second Paley series.

Quasi-Williamson matrices: We recall from [8, Theorem 2.2] the fact that the existence of Ito relative difference sets in the dicyclic group of order $8m$, with m odd, is equivalent to the existence of four generalized Williamson matrices of order m . We coined the name “quasi-Williamson matrices” for this type of generalized Williamson matrices. The four quasi-Williamson matrices have to be circulants but not necessarily symmetric. However, it is required that when plugged into the Williamson array they give a Hadamard matrix of order $4m$. The known series of four Williamson matrices of odd order give rise to the series of NG-pairs. As an example, we have computed four quasi-Williamson matrices of order 35. It is not known whether quasi-Williamson matrices of order 47 exist, and we pose this as an open problem.

Weighing matrices of 2N-type: We apply NG-pairs to the construction of weighing matrices of 2N-type. For small lengths v we list in Appendices B, D and E the NG-pairs of the first and second Paley series and the Ito series, respectively.

k-Toeplitz Hadamard Matrices

We say that a square matrix $\mathbf{A} = [a_{ij}]$, $i, j = 0, 1, \dots, v-1$, is a Toeplitz matrix if $a_{i,j} = a_{i-1,j-1}$ for $i, j > 0$. In particular, we will be interested in two classes of Toeplitz matrices: cyclic (also known as circulant) and negacyclic. The cyclic and negacyclic matrices of order v are polynomials in the cyclic and negacyclic shift matrix $\mathbf{P}$ and $\mathbf{N}$, respectively:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & & 0 & 1 \\ 1 & 0 & 0 & & 0 & 0 \end{bmatrix};$$

$$\mathbf{N} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & & 0 & 1 \\ -1 & 0 & 0 & & 0 & 0 \end{bmatrix}. \quad (1)$$

Definition 1. A k -Toeplitz matrix is a square matrix $\mathbf{A}$ partitioned into square blocks $\mathbf{A}_{ij}$, $i, j = 1, 2, \dots, k$ such that each block $\mathbf{A}_{ij}$ is a Toeplitz matrix. As a special case ($k = 1$), a square Toeplitz matrix is 1-Toeplitz. If each block of a k -Toeplitz matrix is cyclic (resp. negacyclic) we say that it is k -cyclic (resp. k -negacyclic). We abbreviate “ k -Toeplitz”, “ k -cyclic”, “ k -negacyclic” with kT , kC , kN , respectively.

The k -cyclic Hadamard matrices for $k = 1, 2, 4, 8$ have been studied extensively [1, 2, 9–11]. The k -negacyclic ones also have appeared in the literature but to much lesser extent [6, 12]. In this article we are interested mostly in kT -type Hadamard and weighing matrices with $k = 1, 2, 4$.

For $k = 1$ it turns out that Toeplitz Hadamard matrices are necessarily cyclic or negacyclic.

Proposition 1. If $\mathbf{H} = [h_{ij}]$ is a Toeplitz Hadamard matrix of order $v \equiv 0 \pmod{4}$, then $\mathbf{H}$ is cyclic or negacyclic.

Proof: Let $\mathbf{h}_i$ be the $(i + 1)$ th row of $\mathbf{H}$, $\mathbf{h}_i = [h_{i,0}, h_{i,1}, \dots, h_{i,v-1}]$. As the rows of $\mathbf{H}$ are orthogonal to each other, all dot products of two different rows are 0, $\mathbf{h}_i \cdot \mathbf{h}_j = 0$ for $i < j$. Let $j \in \{2, 3, \dots, v-1\}$. Then the equality $\mathbf{h}_0 \cdot \mathbf{h}_{j-1} = \mathbf{h}_1 \cdot \mathbf{h}_j$ simplifies and, by using the hypothesis that $\mathbf{H}$ is a Toeplitz matrix, we deduce that

$$h_{0,v-1}h_{0,v-j} = h_{1,0}h_{j,0}, \quad j = 2, 3, \dots, v-1. \quad (2)$$

Since the entries of $\mathbf{H}$ belong to $\{+1, -1\}$, we have two cases: $h_{1,0} = h_{0,v-1}$ and $h_{1,0} = -h_{0,v-1}$.

In the former case, from the equations (2) we deduce that the equality $h_{j,0} = h_{0,v-j}$ holds for all $j = 1, 2, \dots, v-1$. This means that the matrix $\mathbf{H}$ is cyclic. Similarly, in the latter case one can show that $\mathbf{H}$ is negacyclic.

There is a conjecture, attributed to Ryser [10, p. 134], that there exist no cyclic Hadamard matrices of order > 4 . We conjecture that the negacyclic analog holds.

Conjecture 1. There are no negacyclic Hadamard matrices of order > 2 .

By using a computer we have verified this conjecture for orders ≤ 40 .

For $k = 2$ we shall focus on two special classes of kT -Hadamard matrices, namely the $2C$ - and $2N$ -Hadamard matrices having the form

$$\mathbf{H} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^T & \mathbf{A}^T \end{bmatrix}. \quad (3)$$

($\mathbf{X}^T$ denotes the transpose of a matrix $\mathbf{X}$.)

From now on we refer to 2T-, 2C- and 2N-matrices having the form (3) as *matrices of 2T-type, 2C-type and 2N-type*, respectively.

We propose the following conjecture.

Conjecture 2. *For each even integer $v > 0$ there exists a Hadamard matrix of 2T-type and order $2v$.*

We shall see later that the stronger conjecture below is equivalent to the Ito's conjecture about Hadamard matrices (see [7, 8, 13, 14]).

Conjecture 3. *For each even integer $v > 0$ there exists a Hadamard matrix of 2N-type and order $2v$.*

Three Kinds of Golay Pairs

Let $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$ be a sequence of integers of length v . If each $a_i \in \{\pm 1\}$ then we say that the sequence is *binary*. If we allow the sequence to have also 0s, then we say that it is *ternary*. One defines similarly the binary and ternary matrices. We shall consider $\mathbf{a}$ also as a row-vector.

There are three kinds of autocorrelation functions that we attach to an arbitrary sequence $\mathbf{a}$: the ordinary or nonperiodic (AF), the periodic (PAF), and negaperiodic (NAF) autocorrelation functions. They are defined by the formulas

$$\text{AF}_{\mathbf{a}}(k) = \sum_{i=0}^{v-k-1} a_i a_{i+k}, \quad k \in \mathbf{Z}; \quad (4)$$

$$\text{PAF}_{\mathbf{a}}(k) = \mathbf{a} \cdot \mathbf{a} \mathbf{P}^k, \quad k \in \mathbf{Z}; \quad (5)$$

$$\text{NAF}_{\mathbf{a}}(k) = \mathbf{a} \cdot \mathbf{a} \mathbf{N}^k, \quad k \in \mathbf{Z}, \quad (6)$$

where “ $\cdot$ ” is the dot product. In (4) we use the convention that $a_i = 0$ if $i < 0$ or $i \geq v$.

Note that for $0 \leq k < v$ we have

$$\text{PAF}_{\mathbf{a}}(k) = \text{AF}_{\mathbf{a}}(k) + \text{AF}_{\mathbf{a}}(v-k); \quad (7)$$

$$\text{NAF}_{\mathbf{a}}(k) = \text{AF}_{\mathbf{a}}(k) - \text{AF}_{\mathbf{a}}(v-k). \quad (8)$$

The *cyclic shift* and the *negacyclic shift* of $\mathbf{a}$ are given explicitly by $\mathbf{a} \mathbf{P} = (a_{v-1}, a_0, a_1, \dots, a_{v-2})$ and $\mathbf{a} \mathbf{N} = (-a_{v-1}, a_0, a_1, \dots, a_{v-2})$, respectively.

Since $\mathbf{N}^v = -\mathbf{I}$, we have $\text{NAF}_{\mathbf{a}}(k+v) = -\text{NAF}_{\mathbf{a}}(k)$ for all k . It follows immediately from (8) that

$$\text{NAF}_{\mathbf{a}}(v-k) = -\text{NAF}_{\mathbf{a}}(k), \quad 0 \leq k < v. \quad (9)$$

In particular, if v is even then $\text{NAF}_{\mathbf{a}}(v/2) = 0$. We also mention that $\mathbf{a}$, its reverse sequence and the negashifted sequence $\mathbf{a} \mathbf{N}$ all have the same NAF.

If $\mathbf{A}$ is the negacyclic matrix with first row $\mathbf{a}$, then $\mathbf{A} = \sum_{i=0}^{v-1} a_i \mathbf{N}^i$. Further, $\mathbf{A}^T$ is negacyclic with first row $(a_0, -a_{v-1}, -a_{v-2}, \dots, -a_1)$ and we have

$$\mathbf{A} \mathbf{A}^T = \sum_{k=0}^{v-1} \text{NAF}_{\mathbf{a}}(k) \mathbf{N}^k. \quad (10)$$

(Similar properties are valid for cyclic matrices.)

Let us define three kinds of complementarity:

Definition 2. *The integer sequences $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(t)}$, each of length v , are*

(i) *complementary if $\sum_{i=1}^t \text{AF}_{\mathbf{a}^{(i)}}(k) = 0$ for $k \neq 0$;*

(ii) *P-complementary if $\sum_{i=1}^t \text{PAF}_{\mathbf{a}^{(i)}}(k) = 0$ for $0 < k < v$;*

(iii) *N-complementary if $\sum_{i=1}^t \text{NAF}_{\mathbf{a}^{(i)}}(k) = 0$ for $0 < k < v$.*

We now define three kinds of Golay pairs.

Definition 3. *A Golay pair (G-pair), periodic Golay pair (PG-pair), negaperiodic Golay pair (NG-pair) of length v is a pair $(\mathbf{a}, \mathbf{b})$ of binary sequences of length v which are complementary, P-complementary, N-complementary, respectively. We denote by GP_v , PGP_v and NGP_v the set of Golay, periodic Golay and negaperiodic Golay pairs of length v , respectively.*

For instance, the pair $\mathbf{a} = (1, -1, -1, 1, -1, -1)$, $\mathbf{b} = (1, -1, -1, -1, -1, 1)$ is an NG-pair. It is well known that $\text{GP}_v = \text{PGP}_v = \emptyset$ when v is odd and $v > 1$. We shall see later that this is also true for NGP_v .

The equations (7) and (8) imply that for each $v > 0$ we have $\text{GP}_v = \text{PGP}_v \cap \text{NGP}_v$.

For the definition of equivalence of G-pairs and of PG-pairs see e.g. [9] and [15], respectively. To define the equivalence of NG-pairs $(\mathbf{a}, \mathbf{b})$ of even length v , we introduce the elementary transformations which preserve the set of such pairs:

(i) reverse $\mathbf{a}$ or $\mathbf{b}$;

(ii) replace $\mathbf{a}$ with $\mathbf{a} \mathbf{N}$ or $\mathbf{b}$ with $\mathbf{b} \mathbf{N}$;

(iii) switch $\mathbf{a}$ and $\mathbf{b}$;

(iv) for k relatively prime to v , replace $\mathbf{a}$ and $\mathbf{b}$ with the sequences $(z_i a_{ki(\text{mod } v)})_{i=0}^{v-1}$ and $(z_i b_{ki(\text{mod } v)})_{i=0}^{v-1}$ respectively, where $z_i = 1$ if $ki \pmod{2v} < v$ and $z_i = -1$ otherwise;

(v) replace a_i and b_i with $-a_i$ and $-b_i$, respectively, for each odd index i .

We say that two NG-pairs of the same length are *equivalent* if one can be transformed to the other by a finite sequence of elementary transformations.

As an example, we claim that the NG-pairs $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ of length 10:

$$\mathbf{a} = (+, -, -, -, -, +, -, -, -, -),$$

$$\mathbf{b} = (+, -, -, +, -, +, -, +, +, -);$$

$$\mathbf{c} = (+, -, +, -, +, +, +, -, +, -),$$

$$\mathbf{d} = (+, -, -, +, -, +, -, +, +, -);$$

taken from the Appendices C and D, respectively, are equivalent. (We write “+” and “-” for 1 and -1, respectively.) By applying to $(\mathbf{c}, \mathbf{d})$ the elementary transformation (iv) with $k=9$, we obtain the pair $(\mathbf{a}, \mathbf{d}')$ where $\mathbf{d}' = (+, -, +, -, -, +, +, -, -, +)$.

After reversing $\mathbf{d}'$ and applying the negacyclic shifts, we can transform $\mathbf{d}'$ to $\mathbf{b}$. This proves our claim.

Ito [6] gives a list of NG-pairs of length $v = 2t$ for all odd integers $t \leq 45$. He also points out that no NG-pair of length 94 is known. Apparently this assertion remains still valid.

For lengths $v \leq 40$, the number of equivalence classes in GP_v and their representatives are known (see e.g. [15]). Very recently, such classification has been carried out in [9] for PGP_v with $v \leq 40$.

It is a well-known fact that there is a bijection from PGP_v to the set of 2C-Hadamard matrices of order $2v$. The image of $(\mathbf{a}, \mathbf{b}) \in \text{PGP}_v$ is the matrix (3) in which $\mathbf{a}$ and $\mathbf{b}$ are the first rows of the circulants $\mathbf{A}$ and $\mathbf{B}$. The following is an NG-analog of that result.

Proposition 2. *If $(\mathbf{a}, \mathbf{b})$ is an NG-pair of length v then the matrix (3), where $\mathbf{A}$ and $\mathbf{B}$ are the negacyclic blocks with the first rows $\mathbf{a}$ and $\mathbf{b}$ respectively, is a 2N-type Hadamard matrix of order $2v$. Moreover, this map is a bijection.*

Proof: The formula (10) implies that if $(\mathbf{a}, \mathbf{b}) \in \text{NGP}_v$, then the matrix (3) is a 2N-type Hadamard matrix. The converse also holds.

In view of this proposition we can restate Conjecture 3 as follows:

Conjecture 4. $\text{NGP}_v \neq \emptyset$ for all even $v > 0$.

Let us recall (see [5]) that there are two non-equivalent multiplications

$$\text{GP}_g \times \text{PGP}_v \rightarrow \text{PGP}_{gv}. \quad (11)$$

Interestingly, these two multiplications extend (by using the same formulas) to two multiplications

$$\text{GP}_g \times \text{NGP}_v \rightarrow \text{NGP}_{gv}. \quad (12)$$

Consequently, in order to prove Conjecture 4, it suffices to consider the case when $v \equiv 2 \pmod{4}$.

We can generalize the multiplications (11) and (12) by replacing PG-pairs and NG-pairs with the periodic complementary ternary (PCT) and negaperiodic complementary ternary (NCT) pairs, respectively. We denote by $\text{PCTP}_{v,w}$ and $\text{NCTP}_{v,w}$ the set of PCT-pairs and NCT-pairs of length v and total weight w , respectively. (The weight is the number of nonzero terms.)

Proposition 3. *The Turyn multiplication of Golay pairs (see [11]) extends to maps*

$$\text{GP}_g \times \text{PCTP}_{v,w} \rightarrow \text{PCTP}_{gv,gw}; \quad (13)$$

$$\text{GP}_g \times \text{NCTP}_{v,w} \rightarrow \text{NCTP}_{gv,gw}. \quad (14)$$

Proof: The two proofs are essentially the same and we give the proof only for the case of NCT-pairs. (This proof is similar to the proof of [5, Proposition 3].) Given an integer sequence $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$, we shall represent it by the poly-

nomial $a(z) = a_0 + a_1z + \dots + a_{v-1}z^{v-1}$ in the variable z . The Turyn multiplication $(\mathbf{a}, \mathbf{b})(\mathbf{c}, \mathbf{d}) = (\mathbf{e}, \mathbf{f})$, where $(\mathbf{a}, \mathbf{b}) \in \text{GP}_g$ and $(\mathbf{c}, \mathbf{d}) \in \text{GP}_v$, is given by the formulas

$$e(z) = \frac{1}{2}(a(z) + b(z))c(z^g) + \frac{1}{2}(a(z) - b(z))d(z^{-g})z^{gv-g}; \quad (15)$$

$$f(z) = \frac{1}{2}(b(z) - a(z))c(z^{-g})z^{gv-g} + \frac{1}{2}(a(z) + b(z))d(z^g). \quad (16)$$

The product $(\mathbf{e}, \mathbf{f}) \in \text{GP}_{gv}$.

Now let us assume that $(\mathbf{c}, \mathbf{d}) \in \text{NCTP}_{v,w}$. We define the integer sequences $\mathbf{e}$ and $\mathbf{f}$ of length gv by the same formulas (15) and (16), respectively. It is easy to see that $\mathbf{e}$ and $\mathbf{f}$ are ternary sequences. Since $(\mathbf{a}, \mathbf{b}) \in \text{GP}_g$ we have

$$a(z)a(z^{-1}) + b(z)b(z^{-1}) = 2g. \quad (17)$$

Since $(\mathbf{c}, \mathbf{d}) \in \text{NCTP}_{v,w}$ we have

$$c(z)c(z)^* + d(z)d(z)^* \equiv w \pmod{(z^v + 1)}. \quad (18)$$

This is an identity in the quotient ring $\mathbb{Z}[z]/(z^v + 1)$, which is equipped with the involution “*” sending z to z^{-1} . A computation shows that

$$\begin{aligned} 4e(z)e(z^{-1}) &= (a(z) + b(z))(a(z^{-1}) + b(z^{-1}))c(z^g)c(z^{-g}) + \\ &\quad + (a(z) - b(z))(a(z^{-1}) - b(z^{-1}))d(z^g)d(z^{-g}) + \\ &\quad + (a(z) + b(z))(a(z^{-1}) - b(z^{-1}))c(z^g)d(z^{-g})z^{gv-g} + \\ &\quad + (a(z) - b(z))(a(z^{-1}) + b(z^{-1}))c(z^{-g})d(z^{-g})z^{gv-g}; \\ 4f(z)f(z^{-1}) &= (a(z) - b(z))(a(z^{-1}) - b(z^{-1}))c(z^g)c(z^{-g}) + \\ &\quad + (a(z) + b(z))(a(z^{-1}) + b(z^{-1}))d(z^g)d(z^{-g}) + \\ &\quad + (b(z) - a(z))(a(z^{-1}) + b(z^{-1}))c(z^{-g})d(z^{-g})z^{gv-g} + \\ &\quad + (a(z) + b(z))(b(z^{-1}) - a(z^{-1}))c(z^g)d(z^g)z^{gv-g}. \end{aligned}$$

By using (17) we obtain that

$$e(z)e(z^{-1}) + f(z)f(z^{-1}) = g(c(z^g)c(z^{-g}) + d(z^g)d(z^{-g})).$$

It follows from (18) that

$$c(z^g)c(z^{-g}) + d(z^g)d(z^{-g}) \equiv w \pmod{(z^{gv} + 1)}$$

and so we have

$$e(z)e(z^{-1}) + f(z)f(z^{-1}) \equiv gw \pmod{(z^{gv} + 1)}.$$

We conclude that $(\mathbf{e}, \mathbf{f}) \in \text{NCTP}_{gv,gw}$.

In the special case when $g = 2$ and $(\mathbf{a}, \mathbf{b}) = ((+, -), (+, +))$ we obtain a map $\text{NCTP}_{v,w} \rightarrow \text{NCTP}_{2v,2w}$ to which we refer as “multiplication by 2”.

Cyclic Relative Difference Families

Let us define the map, Φ_v , from the set of binary sequences of length v into the set of v -subsets of the finite cyclic group $\mathbf{Z}_{2v}$ of integers modulo $2v$. If $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$ is a binary sequence then

$$\Phi_v(\mathbf{a}) = \{i : a_i = 1\} \cup \{v + i : a_i = -1\}. \quad (19)$$

Note that Φ_v is injective and that its image consists of all v -subsets $X \subset \mathbf{Z}_{2v}$ such that $i - j \neq v$ for all $i, j \in X$.

We also need the definition of relative difference families in $\mathbf{Z}_{2v}$. They are relative to the subgroup $\{0, v\}$ of order 2.

Definition 4. The subsets $X_1, X_2, \dots, X_s$ of $\mathbf{Z}_{2v}$ form a relative difference family if for each integer $m \in \mathbf{Z}_{2v} \setminus \{0, v\}$ the set of triples $\{(i, j, k) : i, j \in X_k, i - j \equiv m \pmod{2v}\}$ has fixed cardinality λ , independent of m , and there is no such triple if $m = v$.

Note that the parameter λ is uniquely determined by the obvious equation

$$\sum_{i=1}^s k_i (k_i - 1) = 2\lambda(v - 1), \quad (20)$$

where $k_i = |X_i|$ is the cardinality of X_i .

Let us now define the equivalence of relative difference families consisting of two v -subsets $X, Y \subset \mathbf{Z}_{2v}$. First we define five types of elementary transformations which preserve such families:

- (i) replace X or Y with its image by the map $i \rightarrow v - 1 - i \pmod{2v}$;
- (ii) replace X or Y with its image by the map $i \rightarrow i + 1 \pmod{2v}$;
- (iii) switch X and Y ;
- (iv) for k relatively prime to $2v$, replace X and Y with their images by the map $i \rightarrow ki \pmod{2v}$;
- (v) replace X and Y with their images by the map which fixes the even integers and sends $i \rightarrow v + i \pmod{2v}$ if i is odd.

Definition 5. Two relative difference families (X, Y) and (X', Y') on $\mathbf{Z}_{2v}$ are equivalent to each other if one can be transformed to the other by a finite sequence of the above elementary transformations.

Let $(\mathbf{a}, \mathbf{b})$ be a pair of binary sequences of length v and let $X = \Phi_v(\mathbf{a})$ and $Y = \Phi_v(\mathbf{b})$ be the corresponding v -subsets of $\mathbf{Z}_{2v}$. We shall see below that $(\mathbf{a}, \mathbf{b})$ is an NG-pair if and only if (X, Y) is a relative difference family. Moreover, the mapping sending $(\mathbf{a}, \mathbf{b}) \rightarrow (\Phi_v(\mathbf{a}), \Phi_v(\mathbf{b}))$ preserves the equivalence classes. This follows from the fact that Φ_v commutes with the elementary operations (i–v) defined for NG-pairs in the previous section and defined above for relative difference families. For instance, if $\mathbf{a}'$ is the binary sequence obtained from $\mathbf{a}$ by applying the elementary transformation (i), then the set $\Phi_v(\mathbf{a}')$ is obtained from $\Phi_v(\mathbf{a})$ by applying the elementary transformation (i) defined above.

As indicated above, the NG-pairs are closely related to relative difference families. The following two propositions make this more precise.

Proposition 4. Let $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(s)}$ be binary sequences of length v and let $X_1, X_2, \dots, X_s$ be the subsets of $\mathbf{Z}_{2v}$ defined by $X_i = \Phi_v(\mathbf{a}^{(i)})$. If $X_1, X_2, \dots, X_s$ form a relative difference family in $\mathbf{Z}_{2v}$, then the sequences $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(s)}$ are N -complementary.

Proof: We identify the group ring of $\mathbf{Z}_{2v}$ over the integers with the quotient ring $\mathbf{Z}[x]/(x^{2v} - 1)$ of the polynomial ring $\mathbf{Z}[x]$. The cyclic group $\mathbf{Z}_{2v}$ is identified with the multiplicative group $\langle x \rangle$ by the isomorphism sending $i \rightarrow x^i$. The inversion map on $\langle x \rangle$ extends to an involutory automorphism of $\mathbf{Z}[x]/(x^{2v} - 1)$ which we denote by $*$. The subsets X_i are now viewed as subsets of $\langle x \rangle$, and will be identified with the sum of their elements in $\mathbf{Z}[x]/(x^{2v} - 1)$.

Since the X_i form a relative difference family, we have

$$\sum_{i=1}^s X_i X_i^* = \sum_{i=1}^s k_i + \lambda(1 + x^v) \left(x + x^2 + \dots + x^{v-1} \right). \quad (21)$$

The ring of integer negacyclic matrices of order v is isomorphic to the quotient ring $\mathbf{Z}[x]/(x^v + 1)$. It also has an involutory automorphism $*$ which sends x to x^{-1} . Let $f: \mathbf{Z}[x]/(x^{2v} - 1) \rightarrow \mathbf{Z}[x]/(x^v + 1)$ be the canonical homomorphism and note that $f(x^v) = -1$. By applying f to the identity (21) we obtain that

$$\sum_{i=1}^s f(X_i) f(X_i)^* = \sum_{i=1}^s k_i.$$

Note that $f(X_i) = \sum_{j=0}^{v-1} a_j^{(i)} x^j$ and

$$f(X_i) f(X_i)^* = \sum_{j=1}^{v-1} \text{NAF}_{\mathbf{a}^{(i)}}(j) x^j.$$

It follows that $\sum_{j=0}^{v-1} \text{NAF}_{\mathbf{a}^{(i)}}(j) = 0$ for $j = 1, 2, \dots, v - 1$, i.e., the sequences $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(s)}$ are N -complementary.

The following partial converse holds.

Proposition 5. Let $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$ and $\mathbf{b} = (b_0, b_1, \dots, b_{v-1})$ be an NG-pair. Then the subsets $X = \Phi_v(\mathbf{a})$ and $Y = \Phi_v(\mathbf{b})$ form a relative difference family in $\mathbf{Z}_{2v}$ with parameter $\lambda = v$.

Proof: We set $R = \mathbf{Z}[x]/(x^{2v} - 1)$, $R^+ = \mathbf{Z}[x]/(x^v - 1)$ and $R^- = \mathbf{Z}[x]/(x^v + 1)$. Denote the canonical image of $x \in R$ in R^+ and R^- by y and z , respectively. In the proof of Proposition 4 we have defined the involution $*$ in R and R^+ . There is also one in R^- which sends $z \rightarrow z^{-1} = -z^{v-1}$. These involutions commute with the canonical homomorphisms $f: R \rightarrow R^-$ and $g: R \rightarrow R^+$. Note that R is isomorphic to the direct product $R^+ \times R^-$.

Since $(\mathbf{a}, \mathbf{b})$ is an NG-pair, the elements $p, q \in R^-$ defined by $p = \sum a_i z^i$ and $q = \sum b_i z^i$ satisfy

$pp^* + qq^* = 2v$. For convenience we identify X with the sum of its elements in R , and similarly for Y . Then we have $f(X) = p$ and $f(Y) = q$. It follows that $f(XX^* + YY^* - 2v) = 0$. Thus $XX^* + YY^* - 2v$ belongs to the kernel of f and, by using the fact that $(x^v + 1)x^v = x^v + 1$ in R , we obtain an equality

$$\begin{aligned} XX^* + YY^* &= \\ &= 2v + (x^v + 1)(c_0 + c_1x + \dots + c_{v-1}x^{v-1}), \end{aligned} \quad (22)$$

where the c_i are some integers. Since $X = \Phi_v(\mathbf{a})$ and $y^v = 1$, we have $g(X) = 1 + y + \dots + y^{v-1}$. Similarly, $g(Y) = g(X)$. Note that $g(X)^* = g(X)$ and $g(X)^2 = vg(X)$. Hence, by applying g to the equality (22), we obtain that

$$\begin{aligned} 2v(1 + y + \dots + y^{v-1}) &= \\ &= 2v + 2(c_0 + c_1y + \dots + c_{v-1}y^{v-1}). \end{aligned}$$

We deduce that $c_0 = 0$ and $c_i = v$ for $i \neq 0$. The equality (22) now gives

$$XX^* + YY^* = 2v + v(x^v + 1)(x + x^2 + \dots + x^{v-1}).$$

Hence X and Y indeed form a relative difference family in Z_{2v} with the parameter $\lambda = v$.

It was shown in [13, Conjecture 1] that the Ito's conjecture is equivalent to the assertion that for each $t \geq 1$ there exists a relative difference family X_1, X_2 in the cyclic group Z_{4t} with $|X_1| = |X_2| = 2t$ and $\lambda = 2t$. By Propositions 4 and 5 this is in turn equivalent to Conjecture 4.

Paley C-matrices

A *conference matrix* (or *C-matrix*) of order v is a matrix $\mathbf{C}$ of order v whose diagonal entries are 0, the other entries are ± 1 , and such that $\mathbf{C}\mathbf{C}^T = (v-1)\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. There are two well-known necessary conditions for the existence of such matrices. First, v must be even. (We exclude hereafter the trivial case $v = 1$.) Second, if $v \equiv 2 \pmod{4}$ then $v-1$ must be the sum of two squares. For the existence of negacyclic C-matrices of order $v \equiv 4 \pmod{8}$ there is another necessary condition [12], namely that $v-1 = a^2 + 2b^2$ for some integers a and b .

Two C-matrices are said to be *equivalent* if they have the same order and one can be obtained from the other by applying a finite sequence of the following elementary transformations: multiplication of a row or a column by -1 , and interchanging simultaneously two rows and the corresponding two columns.

If $v = 1 + q$ where q is a power of a prime, then Paley [16] has constructed conference matrices of order v . His construction employs essentially the theory of finite fields. Let us recall a general definition as given in [12]. Denote by V a two-dimensional

vector space over the Galois field $\text{GF}(q)$. Choose any set X of $1 + q$ pairwise linearly independent vectors of V . Denote by χ the quadratic character of $\text{GF}(q)$. In particular, $\chi(0) = 0$. (If q is a prime, then χ is the classical Legendre symbol.) Then the matrix

$$\mathbf{C}_X = [\chi(\det(\xi, \eta))], \quad \xi, \eta \in X \quad (23)$$

associated with X , is a C-matrix of order $1 + q$. If $q \equiv 1 \pmod{4}$ then $\chi(-1) = 1$ while when $q \equiv 3 \pmod{4}$ we have $\chi(-1) = -1$. Hence, $\mathbf{C}_X$ is symmetric in the former case and skew-symmetric in the latter case. We refer to $\mathbf{C}_X$ as the *Paley (conference) matrix*. It is known that all Paley conference matrices of the same order are equivalent to each other [17].

In contrast to Conjecture 1, there exist an infinite series of negacyclic C-matrices. Indeed, it is shown in [12, Corollary 7.2] that each Paley C-matrix is equivalent to a negacyclic C-matrix.

Consequently, the following facts hold.

Proposition 6. *Let q be an odd prime power. Then there exist*

- (i) *a negacyclic conference matrix $\mathbf{C}$ of order $1 + q$;*
- (ii) *a $2N$ -type Hadamard matrix $\mathbf{H}$ of order $2(1 + q)$;*
- (iii) *an NG-pair of length $1 + q$.*

Proof: In (ii) we can take $\mathbf{H}$ to be the matrix (3) with $\mathbf{A} = \mathbf{C} + \mathbf{I}$ and $\mathbf{B} = \mathbf{C} - \mathbf{I}$. By Proposition 2, (iii) is equivalent to (ii). Explicitly, if $(0, c_1, c_2, \dots, c_q)$ is the first row of $\mathbf{C}$, then the sequences $(1, c_1, c_2, \dots, c_q)$ and $(-1, c_1, c_2, \dots, c_q)$ form an NG-pair of length $1 + q$.

In Appendix A we list the first rows of the negacyclic Paley C-matrices of order $v = 1 + q \leq 128$.

Let $\mathbf{C}$ be a negacyclic conference matrix of order v with first row $(0, c_1, c_2, \dots, c_{v-1})$. By a theorem of Belevitch [12, Theorem 4.1] we have

$$c_{v/2+j} = (-1)^j c_{v/2-j}, \quad j = 1, 2, \dots, v/2 - 1. \quad (24)$$

One may try to find a counter-example to Conjecture 1 as follows. Let $q \equiv 3 \pmod{4}$ be a prime power. There exists a negacyclic Paley C-matrix $\mathbf{C}$ of order $1 + q$. However, the equations (24) imply that $\mathbf{C}$ is not skew-symmetric. Hence $\mathbf{C} + \mathbf{I}$ is not a Hadamard matrix. On the other hand, we know that $\mathbf{C}$ is equivalent to a skew-symmetric conference matrix $\mathbf{C}'$, and so $\mathbf{C}' + \mathbf{I}$ is a Hadamard matrix. However, $\mathbf{C}' + \mathbf{I}$ is not negacyclic. It appears that $\mathbf{C}$ cannot be used to give a negacyclic Hadamard matrix of order $1 + q$.

The two cases $q \equiv 1 \pmod{4}$ and $q \equiv 3 \pmod{4}$ in Proposition 6 should be considered separately. Indeed, we shall show in a later section that in the latter case the assertion (iii) of Proposition 6 can be made stronger, namely we can replace $1 + q$ by $(1 + q)/2$.

The First Paley Series

We say that any NG-pair $(\mathbf{a}, \mathbf{b})$ of length $v = 1 + q$ resulting from Proposition 6, with $q \equiv 1 \pmod{4}$, belongs to the *first Paley series*. From the proof of that proposition, we recall that $\mathbf{a}$ and $\mathbf{b}$ are the same sequence except that $b_0 = -a_0$.

In this section we assume that q is a prime power and that $q \equiv 1 \pmod{4}$. We recall Theorem 7.3 of [12].

It is easy to verify that if $\mathbf{A}$ is a negacyclic matrix of odd order t and $\mathbf{Z}$ the diagonal matrix of order t with the diagonal elements $1, -1, 1, -1, \dots$, then the matrix $\mathbf{ZAZ}$ is cyclic (and the converse holds).

Proposition 7. *Any Paley conference matrix of order $v = 1 + q \equiv 2 \pmod{4}$, q a prime power, is equivalent to a conference matrix of 2C-type with symmetric circulant blocks.*

Let us give an independent and constructive proof of Proposition 7 in the case of negacyclic conference matrices.

Proof: Let $\mathbf{C}$ be a negacyclic conference matrix of order $v \equiv 2 \pmod{4}$. We shall transform it into the 2N-form, and also into the 2C-form with symmetric blocks.

First, we split the first row $\mathbf{c} = (0, c_1, c_2, \dots, c_{v-1})$ of $\mathbf{C}$ into two pieces $\mathbf{a} = (0, c_2, c_4, \dots, c_{v-2})$ and $\mathbf{b} = (c_1, c_3, \dots, c_{v-1})$. One can easily verify that for each integer k we have $\text{NAF}_{\mathbf{c}}(2k) = \text{NAF}_{\mathbf{a}}(k) + \text{NAF}_{\mathbf{b}}(k)$. It follows that $\mathbf{a}$ and $\mathbf{b}$ are N-complementary sequences. Let $\mathbf{A}$ and $\mathbf{B}$ be the negacyclic matrices with first row $\mathbf{a}$ and $\mathbf{b}$, respectively. By plugging the blocks $\mathbf{A}$ and $\mathbf{B}$ into the array (3), we obtain a C-matrix of 2N-type.

Second, we replace $\mathbf{A}$ and $\mathbf{B}$ with the circulants $\mathbf{ZAZ}$ and $\mathbf{ZBZ}$. The equations (24) imply that the block $\mathbf{ZAZ}$ is symmetric and the first row of $\mathbf{ZBZ}$ is symmetric.

Third, we replace the block $\mathbf{ZBZ}$ with $\mathbf{ZBZP}^m$ where $m = (q-1)/4$. Note that $\mathbf{ZBZP}^m$ is a symmetric circulant. There is no need to change the block $\mathbf{ZAZ}$. By plugging the blocks $\mathbf{ZAZ}$ and $\mathbf{ZBZP}^m$ into the array (3), we obtain a C-matrix of 2C-type with symmetric blocks.

Let us give an example. For $q = 13$ we have $v = 14$ and $m = 3$. From the table in Appendix A, the first row of $\mathbf{C}$ is $\mathbf{c} = (0, +, +, +, +, +, -, -, +, +, -, +)$. Thus, $\mathbf{a} = (0, +, +, -, +, -, -)$ and $\mathbf{b} = (+, +, +, -, +, +, +)$. The first rows of $\mathbf{ZAZ}$ and $\mathbf{ZBZ}$ are $\mathbf{a}' = (0, -, +, +, +, +, -)$ and $\mathbf{b}' = (+, -, +, +, +, -, +)$. Finally, the first row of the circulant $\mathbf{ZBZP}^m$ is $\mathbf{b}'' = (+, +, -, +, +, -, +)$. Thus, the block $\mathbf{ZBZP}^m$ is also symmetric. By plugging the symmetric circulants $\mathbf{A}$ and $\mathbf{B}$ with first rows $\mathbf{a}'$ and $\mathbf{b}''$ into the array (3), we obtain the desired C-matrix of 2C-type.

In Appendix B, for negacyclic Paley C-matrices listed in Appendix A and of order $v \equiv 2 \pmod{4}$, we list the first rows of the symmetric circulant blocks computed by the above procedure.

The Second Paley Series

In this section we denote by $\mathbf{C}$ a negacyclic C-matrix of order $n \equiv 0 \pmod{4}$. For convenience we set $v = n/2$. We give a very simple construction for NG-pairs of length v . In particular we can take $n = 1 + q$ where $q \equiv 3 \pmod{4}$ is a prime power. Indeed, as mentioned earlier, we know that any Paley C-matrix of order $1 + q$ is equivalent to a negacyclic C-matrix. We point out that we do not have any other examples of matrices $\mathbf{C}$.

Proposition 8. *Let $\mathbf{C}$ be a negacyclic C-matrix of order $n \equiv 0 \pmod{4}$. If $\mathbf{c} = (0, c_1, c_2, \dots, c_{n-1})$ is the first row of $\mathbf{C}$, then the sequences $\mathbf{a} = (1, c_2, c_4, \dots, c_{n-2})$ and $\mathbf{b} = (c_1, c_3, \dots, c_{n-1})$ form an NG-pair of length $v = n/2$.*

Proof: For convenience, we set $\mathbf{a}' = (0, c_2, c_4, \dots, c_{n-2})$. Then $\text{NAF}_{\mathbf{a}'}(k) + \text{NAF}_{\mathbf{b}}(k) = \text{NAF}_{\mathbf{c}}(2k)$ for $k = 1, 2, \dots, v-1$. Since $\mathbf{C}$ is a conference matrix, it follows from (10) that $\text{NAF}_{\mathbf{c}}(k) = 0$ for $k = 1, 2, \dots, n-1$. Hence, $(\mathbf{a}', \mathbf{b})$ is an N-complementary pair. However, this is not an NG-pair because the first term of $\mathbf{a}'$ is 0.

Let us write $\mathbf{a}'' = (x, a_1, a_2, \dots, a_{v-1})$ with $a_i = c_{2i}$ for $i = 1, 2, \dots, v-1$ and x an integer variable. We claim that $\text{NAF}_{\mathbf{a}''}(k) = \text{NAF}_{\mathbf{a}'}(k)$ for $0 < k < v$. Indeed, we have $\text{NAF}_{\mathbf{a}''}(k) = \text{AF}_{\mathbf{a}''}(k) - \text{AF}_{\mathbf{a}''}(v-k) = \text{NAF}_{\mathbf{a}'}(k) + x(a_k - a_{v-k})$. By Belevitch's theorem, we have $a_k = a_{v-k}$ for $0 < k < v$ and so $\text{NAF}_{\mathbf{a}''}(k) = \text{NAF}_{\mathbf{a}'}(k)$. Thus our claim is proved.

If we now set $x = 1$ then $\mathbf{a}'' = \mathbf{a}$ and we conclude that $\text{NAF}_{\mathbf{a}}(k) = \text{NAF}_{\mathbf{a}'}(k)$ for $0 < k < v$. Consequently, $(\mathbf{a}, \mathbf{b})$ is an NG-pair.

We say that the NG-pairs constructed in this proposition belong to the *second Paley series*. We say that an NG-pair is a *Paley NG-pair* if it belongs to the first or the second Paley series.

Out of the 63 odd positive integers $t \leq 125$, there are exactly 18 for which there is no Paley NG-pair of length $v = 2t$. Let us list these integers:

$$23, 29, 39, 43, 47, 59, 65, 67, 73, 81, \\ 89, 93, 101, 103, 107, 109, 113, 119. \quad (25)$$

In Appendix C we list the NG-pairs in the second Paley series obtained from the negacyclic C-matrices listed in Appendix A with $q \equiv 3 \pmod{4}$.

Ito Series

There is another series, due to Ito [7], of NG-pairs of length $(1+q)/2$ when $q \equiv 3 \pmod{4}$ is a prime power. However, we will show below that the NG-pairs in this series belong to the second Paley series.

For convenience we set $t = (1+q)/4 = v/2$ and let p be the prime such that $q = p^n$. The Ito series is derived from the relative difference sets constructed

by Ito [7]. These relative difference sets R have parameters $(4t, 2, 4t, 2t)$ and lie in the dicyclic group

$$\text{Dic}_{8t} = \langle a^{4t} = 1, b^2 = a^{2t}, bab^{-1} = a^{-1} \rangle \quad (26)$$

of order $8t$. The forbidden subgroup is $\langle b^2 \rangle$.

For convenience we identify a subset $X \subseteq \text{Dic}_{8t}$ with the sum of its elements in the group-ring (over $\mathbb{Z}$) of Dic_{8t} . Then we can write $R = R_1 + R_2b$ with $R_1, R_2 \subseteq \langle a \rangle$. The sets R_1 and R_2 form a relative difference family in the cyclic group $\langle a \rangle$ (with the same forbidden subgroup). Let us identify $\langle a \rangle$ with $\mathbb{Z}_{4t}$ by the isomorphism sending $a \rightarrow 1$. It is obvious that R_1 and R_2 are $2t$ -subsets of $\mathbb{Z}_{4t}$. By Proposition 4, the binary sequences $X_1 = \Phi_v^{-1}(R_1)$ and $X_2 = \Phi_v^{-1}(R_2)$ form an NG-pair.

We shall now describe a procedure which takes as input the integer t and a primitive polynomial f of degree $2n$ over the prime field $\text{GF}(p) = \mathbb{Z}_p$, and gives as output the NG-pair arising from the Ito's difference set R in Dic_{8t} . This procedure is based on the simplification of Ito's construction due to B. Schmidt [8, Theorem 3.3].

We construct the Galois field $\text{GF}(q^2)$ by adjoining a root x of f to $\mathbb{Z}_p$. As $q^2 - 1 = ((q - 1)/2)(2q + 1)$ and $(q - 1)/2 = 2t - 1$ and $2q + 1 = 8t$ are relatively prime, the multiplicative group $\text{GF}(q^2)^*$ is a direct product of the subgroups U of order $(q - 1)/2$ and W of order $2q + 1$. Note that U is the subgroup of squares in $\text{GF}(q^2)^*$. (Thus we have $Q = U$ for the set Q defined in the proof of [8, Theorem 3.3].)

As f is primitive, x generates $\text{GF}(q^2)^*$ and the elements $u = x^{8t}$ and $w = x^{2t-1}$ generate U and W , respectively. Since $x^{(q^2-1)/2} = -1$, the element $\alpha = x^{2t}$ satisfies the equation $\alpha + \alpha^q = 0$, i.e., $\text{tr}(\alpha) = 0$ where $\text{tr}: \text{GF}(q^2) \rightarrow \text{GF}(q)$ is the (relative) trace map. We set $v = 2t$ and define two binary sequences $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$ and $\mathbf{b} = (b_0, b_1, \dots, b_{v-1})$ of length v . We declare that $a_i = 1$ if and only if $\text{tr}(\alpha w^{2i}) \in U$, and declare that $b_i = 1$ if and only if $\text{tr}(\alpha w^{2i+1}) \in U$. Then $(\mathbf{a}, \mathbf{b}) \in \text{NGP}_v$. Note that $a_0 = -1$.

We say that the NG-pairs obtained by this procedure belong to the *Ito series*. They exist for lengths $v = 2t$ where $q = 4t - 1$ is a prime power.

For a sequence $\mathbf{a} = (a_0, a_1, \dots, a_{v-1})$ we say that it is *quasi-symmetric* if $a_i = a_{v-i}$ for $i = 1, 2, \dots, v - 1$. Note that the negacyclic matrix with first row $\mathbf{a}$ is skew-symmetric if and only if $\mathbf{a}$ is quasi-symmetric and $a_0 = 0$.

The Ito NG-pairs $(\mathbf{a}, \mathbf{b})$ have some additional symmetries. Namely, $\mathbf{a}$ is quasi-symmetric and $\mathbf{b}$ is skew-symmetric. Both assertions follow from the fact that

$$\begin{aligned} \text{tr}(\alpha w^{8t-i}) &= \text{tr}(\alpha w^{-i}) = \alpha(w^{-i} - w^{-iq}) = \\ &= \alpha w^{-i(q+1)}(w^{iq} - w^i) = (-1)^i \text{tr}(\alpha w^i). \end{aligned}$$

These symmetry properties were observed by Ito [7, Proposition 6], as well as the fact that the

2N-type Hadamard matrix constructed from the NG-pair $(-\mathbf{a}, \mathbf{b})$ is skew-Hadamard. (Since the diagonal entries of a skew-Hadamard matrix have to be equal to $+1$, we replaced $\mathbf{a}$ with $-\mathbf{a}$.)

It follows from these symmetry properties that the negacyclic matrix with first row

$$(0, b_0, a_1, b_1, \dots, a_{v-1}, b_{v-1})$$

is a conference matrix. This shows that the NG-pair $(-\mathbf{a}, \mathbf{b})$ belongs to the second Paley series.

In Appendix D we list the NG-pairs of length $v = (1 + q)/2 \leq 154$ in the Ito series, with $q \equiv 3 \pmod{4}$ a prime power. We have verified directly that each NG-pair listed in Appendix C is equivalent to the corresponding NG-pair (the one having the same length, v) in the list of Appendix D.

There exist prime powers $q > 1$ such that $q \equiv 1 \pmod{4}$ and $1 + 2q$ is also a prime power. For instance, $q = 5, 9, 13, 29, 41$. For such q there exist NG-pairs $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ of length $1 + q$ which belong to the first and the second Paley series, respectively. Then the following question arises: can $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ be equivalent? (We believe that the answer is negative.)

Quasi-Williamson Matrices

We say that four binary matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ of order t are *quasi-Williamson matrices* if they are circulants and satisfy the equations

$$\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T + \mathbf{C}\mathbf{C}^T + \mathbf{D}\mathbf{D}^T = 4t\mathbf{I}; \quad (27)$$

$$\mathbf{A}\mathbf{B}^T + \mathbf{C}\mathbf{D}^T = \mathbf{B}\mathbf{A}^T + \mathbf{D}\mathbf{C}^T. \quad (28)$$

This is the cyclic case of a more general definition given in [8]. In order to avoid a possible confusion, we have introduced a different name for this type of matrices. Note that the above two equations amount to saying that the matrix

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \\ -\mathbf{B} & \mathbf{A} & -\mathbf{D} & \mathbf{C} \\ -\mathbf{C}^T & \mathbf{D}^T & \mathbf{A}^T & -\mathbf{B}^T \\ -\mathbf{D}^T & -\mathbf{C}^T & \mathbf{B}^T & \mathbf{A}^T \end{bmatrix} \quad (29)$$

is a Hadamard matrix.

The *Williamson matrices* are the special case of quasi-Williamson matrices where we require all four blocks $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ to be symmetric, in which case the condition (28) is automatically satisfied. Let us mention the following two infinite series of Williamson matrices of order t . The first, due to Turyn, exists in orders $t = (1 + q)/2$, where $q \equiv 1 \pmod{4}$ is a prime power. Given a conference matrix of 2C-type, see Proposition 7, with symmetric circulant blocks, say $\mathbf{A}$ and $\mathbf{B}$, then the matrices $\mathbf{A} + \mathbf{I}$, $\mathbf{A} - \mathbf{I}$, $\mathbf{B}$, $\mathbf{B}$ are four Williamson matrices (this is the

Turyn series). The second, due to Whiteman, exists in orders $t = p(1 + p)/2$, where $p \equiv 1 \pmod{4}$ is a prime.

In the rest of this section we assume that t is odd. Then quasi-Williamson matrices of order t are equivalent to relative difference sets in Dic_{8t} [8].

Let $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ be the first rows of quasi-Williamson matrices of order t . We set $z = 1$ if $t \equiv 1 \pmod{4}$ and $z = -1$ otherwise. We shall describe a procedure which takes as input the quadruple $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ and gives as output an NG-pair of length $v = 2t$. It is based on the proof of [8, Theorem 2.1]. The subgroup $G \times \langle x \rangle$ of the group $G \times Q_8$, in the mentioned proof, is cyclic and is identified with $\mathbf{Z}_{4t}$.

By using the rows $\mathbf{a}$ and $\mathbf{b}$, we construct a binary sequence p of length v as follows. Say, $\mathbf{a} = (a_0, a_1, \dots, a_{t-1})$. We define two subsets $\mathbf{a}'$, $\mathbf{a}''$ of $\mathbf{Z}_t$ by $\mathbf{a}' = \{i : a_i = 1\}$ and $\mathbf{a}'' = \{i : a_i = -1\}$. We define similarly the subsets $\mathbf{b}'$, $\mathbf{b}'' \subseteq \mathbf{Z}_t$.

For $i = 0, 1, 2, 3$ we define the map $\Psi_i: \mathbf{Z}_t \rightarrow \mathbf{Z}_{4t}$ by the formula

$$\Psi_i(j) = j + t(z(i-j) \pmod{4}). \quad (30)$$

It is easy to verify that the set

$$X = \Psi_0(\mathbf{a}') \cup \Psi_1(\mathbf{b}') \cup \Psi_2(\mathbf{a}'') \cup \Psi_3(\mathbf{b}'')$$

lies in the image of the map Φ_v [see (19)]. Finally, we set $p = \Phi_v^{-1}(X)$, which is a binary sequence of length v .

Similarly, from **c** and **d** we construct first a v -subset $Y \subseteq \mathbf{Z}_{4t}$ and then the binary sequence $q = \Phi_v^{-1}(Y)$, of length v . Then $(p, q) \in \text{NGP}_v$.

We remark that the v -subsets X and Y form a relative difference family in $\mathbf{Z}_{2v}$ with parameter $\lambda = v$ and the forbidden subgroup $\{0, v\}$.

The converse is also true: given an NG-pair $(\mathbf{a}, \mathbf{b})$ of length $2t$ we can construct quasi-Williamson matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ of order t . As an example, we used the NG-pair of length $v = 70$ given in Appendix D to compute four quasi-Williamson matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ of order 35. The first rows of these matrices (after some cyclic shifts) are:

$$\mathbf{a} = [+,-,+,+,-,+,-,+,+,+,+,-,+,+,+,-, \\ -,+,+,-,-,-,+,+,-,-,-,-,+,+,-,-,+],$$

$$\mathbf{b} = [+ , + , + , + , + , - , - , - , + , - , + , + , - , + , + , + , - , + , + , - , + , + , + , - , + , + , - , + , - , - , - , + , + , + , +];$$

$$\mathbf{c} = [-, +, -, +, +, -, +, -, +, +, +, -, -, -, -, +, +, +, \\ +, +, +, -, -, +, +, -, +, -, +, -, -, -, +, +, -],$$

$$\mathbf{d} = [-, +, +, -, -, -, +, -, +, -, +, +, -, -, +, +, +, +, +, +, -, -, -, -, +, +, +, -, +, -, +, +, -, +, -].$$

respectively. The blocks **A**, **B**, **C**, **D** satisfy the equations (27) and (28), and when plugged into the array (29) we do get a Hadamard matrix. Moreover, **A** is of skew-type, while **B** is symmetric, and **d** is the

reverse of $\mathbf{c}$. Note also that $\mathbf{A}\mathbf{B}^T - \mathbf{B}\mathbf{A}^T = (\mathbf{A} - \mathbf{A}^T)\mathbf{B} \neq \mathbf{0}$, and so $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ are not matrices of Williamson type according to [2, Definition 3.3].

It is known that Williamson matrices of odd order t exist for $t = 23, 29, 39, 43$, see e.g. [18]. After removing these integers, the list (25) reduces to

$$47, 59, 65, 67, 73, 81, 89, 93, 101, \\ 103, 107, 109, 113, 119. \quad (31)$$

Let us single out the smallest case.

Open Problem. Do quasi-Williamson matrices of order 47 exist? Equivalently, do NG-pairs of length 94 exist?

The above mentioned facts have been known since 1999 [8, 6] and apparently no progress has been made so far in the search for NG-pairs of order $v = 2t$, for t in the above list. For generalizations where the cyclic group $\mathbf{Z}_{4t}$ is replaced by more general finite abelian groups see [14].

Since the known infinite series of NG-pairs are rather sparse, it is hard to believe that NG-pairs exist for all even lengths. In other words, in our opinion Ito's conjecture is likely to be false.

Weighing Matrices of 2N-type

A *weighing matrix* of order n and *weight* w (abbreviated as $\mathbf{W}(n, w)$) is a matrix $\mathbf{W}$ of order n with entries in $\{0, \pm 1\}$ such that $\mathbf{W}\mathbf{W}^T = w\mathbf{I}$. In this section we discuss the existence of weighing matrices of 2N-type.

Note that C-matrices of order v are $\mathbf{W}(v, v-1)$. It is known that there are no cyclic $\mathbf{W}(v, v-1)$ for $v > 2$ [19]. On the other hand there are infinitely many negacyclic $\mathbf{W}(v, v-1)$. Indeed each Paley C-matrix is equivalent to a negacyclic C-matrix. It has been conjectured [12] that there are no negacyclic C-matrices of even order $v \neq 1 + q$, q a prime power. This conjecture has been verified for $v \leq 226$. However, there exist C-matrices of 2N-type whose order v is not of that form. For instance, they exist for

$$v = 16, 40, 52, 56, 64, 88, 96, 120, 136, 144, 160.$$

(See part (iii) of the proposition below.)

We have four infinite series of 2N-type weighing matrices.

Proposition 9. *Let q be an odd prime power. Then there exist weighing matrices of $2N$ -type:*

- (i) $\mathbf{W}(1+q, q)$;
- (ii) $\mathbf{W}(2+2q, 2q)$;
- (iii) if $q \equiv 3 \pmod{4}$, $\mathbf{W}(2+2q, 1+2q)$ and $\mathbf{W}(4+4q, 2+4q)$.

Proof: (i) If $q \equiv 1 \pmod{4}$, this was shown in the proof of Proposition 7. Otherwise the claim follows from the fact (proven in the Ito series section), that there exists an NG-pair $(\mathbf{a}, \mathbf{b})$ of length $(1+q)/2$ with $\mathbf{a}$ quasi-symmetric. Let $\mathbf{A}$ and $\mathbf{B}$ be the negacyclic

matrices with first rows **a** and **b**. We may assume that $a_0 = 1$, then the matrix (3) is skew-Hadamard of 2N-type. By replacing the diagonal entries with 0s, we obtain a $\mathbf{W}(1 + q, q)$.

(ii) This follows from (i) because we can “multiply by 2”.

(iii) Let **(a, b)** be an Ito NG-pair of length $(1 + q)/2$. By multiplying by 2, we obtain an NG-pair **(a', b')** of length $1 + q$ with $\mathbf{a}' = (1, \mathbf{a}'')$ quasi-symmetric. Consequently, the pair **((0, a''), b')** is N-complementary. The corresponding 2N-type matrix (3) is a C-matrix of order $2 + 2q$. Multiplying by 2 we obtain also an $\mathbf{W}(4 + 4q, 2 + 4q)$.

This proposition covers all weighing matrices $\mathbf{W}(4n, 4n - 1)$ and $\mathbf{W}(4n, 4n - 2)$ of 2N-type, for $n \leq 50$ except for

$$n = 9, 13, 19, 23, 25, 28, 29, 31, 37, 39, 43, 44, \\ 46, 47, 48, 49$$

and

$$n = 11, 17, 18, 26, 29, 33, 35, 38, 39, 43, 46, 47, 50,$$

respectively. We have constructed five of these matrices:

n **a & b**

$$11 \quad [0, -, -, +, -, -, -, -, +, +, +, -, +, +, -, +, -, +, \\ +, +, -], [0, +, -, -, -, -, +, -, -, +, -, +, +, +, -, -, \\ -, -, +, -, +, +]$$

$$13 \quad [0, +, +, -, +, -, +, +, +, +, -, +, +, -, +, +, \\ +, +, -, +, -, +, +], [+ , +, +, -, +, -, -, -, +, -, -, \\ -, +, +, +, -, +, +, -, +, +, +, +, -, -] \\ 17 \quad [0, -, -, -, +, -, -, +, -, +, -, -, -, +, -, +, +, +, -, \\ +, +, +, +, -, +, +, +, +, +, -, -, -, +, -], [0, -, +, \\ +, +, +, -, +, -, -, -, +, +, +, +, -, -, +, +, -, -, +, \\ -, +, +, -, +, +, +, +, -, +, -, -] \\ 18 \quad [0, +, +, -, -, -, +, -, -, +, -, -, -, -, -, +, +, \\ +, -, +, +, -, +, -, +, -, -, -, +, +, +, -, +, +, -], \\ [0, +, +, -, -, +, +, +, +, -, +, -, -, -, +, -, +, +, +, \\ -, +, +, +, +, -, +, +, +, +, -, +, -, -, +, +, -] \\ 26 \quad [0, +, +, +, +, -, -, -, +, -, -, -, +, +, +, +, +, -, +, \\ -, -, +, +, -, +, -, -, -, +, +, +, +, -, +, -, +, +, \\ +, +, +, +, -, -, +, +, -, -, +, -, -, +, +, -, -, +, +, \\ +, +, -, -, +, +, -, -, +, -, -, +, +, +, +, -, -, +, +, \\ -, +, +, -, +, +, +, +, -, -, +, +, +, +, -, -, +, +, \\ -, +, +, +, -, +, +, -, +, -].$$

Multiplication by Golay pairs may be used to construct other series of weighing matrices of 2N-type.

In Appendix E we list weighing matrices $\mathbf{W}(4n, 4n - 2)$ of 4C-type for odd $n \leq 21$. They can be easily converted to 4N-type by replacing each circulant block **X** of order n with the negacyclic block $\mathbf{ZXZ}$.

Acknowledgements: The second author wishes to acknowledge generous support by NSERC. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET) and Compute/Calcul Canada.

Appendices

Appendix A. For even integers $v = 1 + q \leq 128$, with $q = p^n$ a power of a prime p , we give the first row **c** of a negacyclic conference matrix **C** of order v belonging to the equivalence class of Paley conference matrices. The algorithm is described in the section on Paley C-matrices, it is based on [12, Corollary 7.2]. We also record the primitive polynomial $f(x)$ of degree $2n$ over $\text{GF}(p)$ used in the computation.

v $f(x); p, q$ and the first row **c**

$$4 \quad x^2 + x + 2; p = q = 3 \\ [0, +, -, -] \\ 6 \quad x^2 + x + 2; p = q = 5 \\ [0, +, +, +, -, +] \\ 8 \quad x^2 + x + 3; p = q = 7 \\ [0, +, -, -, -, +, -, -] \\ 10 \quad x^4 + x^3 + 2; p = 3, q = 9 \\ [0, +, -, -, -, -, +, -, +, +] \\ 12 \quad x^2 + x + 7; p = q = 11 \\ [0, +, -, +, -, -, +, +, -, -, -, -] \\ 14 \quad x^2 + x + 2; p = q = 13 \\ [0, +, +, +, +, -, -, +, +, -, +, -, +] \\ 18 \quad x^2 + x + 3; p = q = 17 \\ [0, +, +, +, -, +, +, +, +, -, -, +, -, +, +, -, +] \\ 20 \quad x^2 + x + 2; p = q = 19 \\ [0, +, -, -, -, -, +, -, -, +, +, -, -, -, +, -, +, -, -]$$

- [illegible]

- [illegible]

Appendix B. For even integers $v = 1 + q \leq 128$, with $q \equiv 1 \pmod{4}$ a prime power, we give a 2C-type conference matrix of order v with symmetric blocks **A** and **B** which belongs to the equivalence class of Paley conference matrices. The algorithm is described in the section on the first Paley series. Since the blocks **A** and **B** are symmetric circulants of odd order $v/2$, we record only the first $(v + 2)/4$ elements of their first rows **a** and **b**.

We recall that $\mathbf{A} + \mathbf{I}$, $\mathbf{A} - \mathbf{I}$, $\mathbf{B}$, $\mathbf{B}$ are four Williamson matrices of order $v/2$ belonging to the Turyn series.

v first rows of **a** and **b** (truncated)

- 6 [0, −], [−, +]
10 [0, +, −], [−, +, +]
14 [0, −, +, +], [+ , +, −, +]
18 [0, −, −, −, +], [−, −, +, −, +]
26 [0, +, −, −, −, +, −], [−, +, −, −, +, +, +]
30 [0, −, +, −, +, +, −, −], [−, −, +, +, −, +, +, +]
38 [0, −, +, +, −, −, +, −, −, −], [−, −, −, +, +, +, −, +, −, +]
42 [0, −, +, −, +, +, −, −, +, +, +], [−, +, −, −, −, −, +, −, −, +, +]
50 [0, +, −, −, +, −, −, −, +, +, +, +, −], [−, +, +, −, −, +, −, +, −, +, +, +, +]
54 [0, −, +, +, −, +, −, −, −, +, −, −, +, +], [+ , +, +, +, −, +, +, +, −, −, −, +, −, +]
62 [0, −, +, −, +, +, +, +, −, +, −, −, +, +, +, −], [−, −, +, −, −, −, +, +, +, +, +, −, +, +, −, +]
74 [0, −, −, +, +, +, −, +, −, −, −, +, −, −, −, −, +, +, −], [+ , +, −, −, −, −, +, −, −, +, −, −, +, −, +, +, +]
82 [0, −, −, +, +, +, +, −, +, −, +, −, −, +, +, −, +, −, −, +],
[−, −, +, −, −, −, +, +, −, −, −, −, −, +, −, −, +, −, +, +]
90 [0, −, +, −, +, −, −, −, −, +, +, +, +, +, +, +, −, +, +, −, +, +, −],
[+ , −, +, +, −, −, −, −, +, +, +, −, +, −, +, +, −, +, +, +, −, −, +]
98 [0, −, +, +, +, −, −, +, +, +, −, +, −, −, +, +, +, −, +, −, −, −, −, −],
[+ , −, +, +, +, +, −, −, +, +, +, −, +, +, +, +, +, +, −, −, −, +, +]
102 [0, −, +, +, −, −, −, +, +, −, +, −, −, −, +, +, −, +, +, −, +, −, −, −, −],
[−, −, −, +, −, −, +, +, −, −, −, +, +, +, −, +, −, +, +, −, −, +, +, +]

- 110 [0, -, -, +, -, -, -, +, -, -, -, +, +, +, -, -, +, -, +, +, +, -, -, +, -, -],
 [-, +, +, -, +, +, -, +, +, +, -, -, -, -, -, +, +, -, +, -, +, -, -, -, +]
 114 [0, -, -, +, +, +, +, -, -, +, -, +, -, -, +, +, +, -, +, +, -, -, -, -, -, -, -],
 [+ , +, -, +, -, -, -, +, -, +, -, -, +, -, +, +, -, -, +, +, +, -, -, -, -, +, -, -, +]
 122 [0, +, -, +, -, -, +, +, +, +, +, -, -, -, +, +, -, +, -, +, +, +, +, -, -, +, -, -, -, -, -, -],
 [-, +, -, -, +, +, +, -, +, +, -, +, -, -, +, +, -, +, +, +, -, -, -, +, -, +, +, +, -, +, +]

Appendix C. For integers $q = 4t - 1$, with $q = p^n \equiv 3 \pmod{4}$ a power of a prime p , we give the NG-pairs $(\mathbf{a}, \mathbf{b})$ of length $v = 2t \leq 64$ belonging to the second Paley series. The procedure used to generate this list is described in the section on the second Paley series.

The sequence $\mathbf{a}$ is quasi-symmetric and $\mathbf{b}$ is skew-symmetric. We record only the first $t + 1$ terms of $\mathbf{a}$ and the first t terms of $\mathbf{b}$. If $\mathbf{A}$ and $\mathbf{B}$ are the negacyclic blocks with first rows $\mathbf{a}$ and $\mathbf{b}$, then the matrix (3) is $2N$ -type skew-Hadamard.

v $\mathbf{a} \ \& \ \mathbf{b}$ (truncated)

- 2 [+ , -], [+]
 4 [+ , - , -], [+ , -]
 6 [+ , - , -, +], [+ , + , -]
 10 [+ , -, -, -, -, +], [+ , -, -, + , -]
 12 [+ , -, + , + , + , -, +], [+ , -, + , + , + , +]
 14 [+ , + , -, -, + , + , + , +], [+ , -, + , -, + , + , +]
 16 [+ , -, -, + , + , -, + , -, -], [+ , + , + , + , -, -, + , -]
 22 [+ , -, -, -, + , -, -, + , + , -, + , -], [+ , + , + , + , -, + , -, -, + , + , +]
 24 [+ , -, + , + , -, + , -, -, -, + , + , + , -], [+ , -, + , -, + , + , + , + , + , -, -]
 30 [+ , -, + , -, + , -, -, -, + , -, -, -, -, + , +], [+ , -, -, + , -, -, -, + , + , + , + , -, + , + , -]
 34 [+ , -, + , + , + , -, + , -, + , -, -, -, -, -, + , +], [+ , + , -, + , + , + , + , + , -, -, -, + , -, -, + , + , -]
 36 [+ , -, -, -, -, -, + , + , -, + , -, + , + , + , -, + , + , +], [+ , -, + , -, -, + , + , -, + , -, -, -, + , + , + , -, +]
 40 [+ , -, + , -, -, -, -, + , -, + , -, -, -, -, -, + , + , -], [+ , -, + , + , + , + , -, -, -, -, -, + , + , -, -, + , + , -, + , -]
 42 [+ , -, + , + , + , + , + , -, -, + , -, + , + , + , + , -, + , + , -, -, + , +],
 [+ , -, -, + , -, + , -, + , -, + , + , -, -, -, + , -, -, -, -]
 52 [+ , -, + , -, -, -, -, -, + , -, + , + , + , + , -, -, -, + , + , -, + , + , -, + , + , -],
 [+ , -, -, + , + , + , -, -, -, -, -, + , -, -, -, -, + , -, + , -, + , -, -, + , + , -]
 54 [+ , -, + , -, -, -, -, -, + , + , -, + , + , -, + , -, -, -, + , -, + , -, + , -, -, -],
 [+ , -, + , + , + , -, -, + , + , -, -, -, -, -, -, + , + , -, + , -, -, + , + , -]
 64 [+ , -, -, -, + , -, + , -, -, -, -, -, + , + , -, + , + , -, -, + , + , + , + , + , + , -, + , + , + , -, -],
 [+ , -, + , + , -, -, + , -, -, -, -, -, + , + , + , + , -, + , -, + , + , -, + , -, + , + , + , -, -, -, + , + , -]

Appendix D. For integers $q = 4t - 1$, with $q = p^n \equiv 3 \pmod{4}$ a power of a prime p , we give the NG-pairs $(\mathbf{a}, \mathbf{b})$ of length $v = 2t \leq 154$ belonging to the Ito series. The procedure used to generate this list is described in the Ito series section. In the list below, for each length v , we record the primitive polynomial $f(x)$ of degree $2n$ over $\text{GF}(p)$ used in the computation, and the NG-pair $(\mathbf{a}, \mathbf{b})$.

In all cases we have $\mathbf{a} = (+, \mathbf{a}')$ where the subsequence $\mathbf{a}'$ is symmetric while the whole sequence $\mathbf{b}$ is skew-symmetric. We record only the first $t + 1$ terms of $\mathbf{a}$ and the first t terms of $\mathbf{b}$. If $\mathbf{A}$ and $\mathbf{B}$ are the negacyclic blocks with first rows $\mathbf{a}$ and $\mathbf{b}$, then the matrix (3) is skew-Hadamard of $2N$ -type.

Moreover, by multiplying the NG-pair $(\mathbf{a}, \mathbf{b})$ by 2, we obtain in the same way a $2N$ -type skew-Hadamard matrix of order $1 + q$.

v $\mathbf{a} \ \& \ \mathbf{b}$ (truncated)

- 2 $x^2 - x - 1; p = q = 3$
 [+ , +], [+]
 4 $x^2 - x + 3; p = q = 7$
 [+ , -, +], [+ , +]
 6 $x^2 + x + 7; p = q = 11$
 [+ , -, + , +], [- , -, +]
 10 $x^2 - x + 2; p = q = 19$
 [+ , -, + , -, + , +], [+ , + , + , -, -]

- [illegible]

Appendix E. We list here the weighing matrices $W(4n, 4n-2)$ of 4C-type for odd $n \leq 21$.

$4n$ **a, b, c, d**

4	[0], [+], [0], [+]
12	[0, +, +], [+ , - , -], [0, - , -], [+ , - , -]
20	[0, +, +, +, +], [+ , + , - , - , +], [0, +, - , - , +], [+ , - , +, +, -]
28	[0, - , +, +, +, +, -], [+ , + , - , +, +, - , +], [0, - , +, - , - , +, -], [+ , + , +, - , - , +, +]
36	[0, +, - , +, - , - , +, - , +], [+ , + , - , - , - , - , - , +], [0, - , +, +, - , - , +, +, -], [+ , + , +, - , +, +, - , +, +]
44	[0, +, - , - , +, - , - , +, - , +], [+ , + , - , - , - , - , - , - , +], [0, +, - , +, - , +, +, - , +, - , +], [+ , + , +, - , - , +, +, - , - , +, +]
52	[0, +, +, - , +, - , - , - , +, +, +], [+ , - , +, - , - , - , +, +, - , - , +, -], [0, - , - , +, +, +, - , - , +, +, +, -], [+ , - , - , +, - , - , - , - , - , +, - , -]
60	[0, - , - , +, +, +, - , +, +, - , +, +, +, -], [+ , + , - , - , - , - , +, - , - , - , - , +], [0, +, +, - , +, - , - , - , +, +, - , +], [+ , - , +, +, - , - , - , - , - , - , +, +, -]
68	[0, - , +, +, +, - , +, - , - , - , +, - , +, +, -], [- , +, - , - , +, - , +, +, - , +, - , - , +], [0, +, +, +, +, - , - , +, +, - , - , +, +, +, +], [+ , + , +, - , +, - , - , - , +, +, - , - , +, +]
76	[0, +, +, +, +, - , +, - , - , +, +, - , +, +, +, +], [- , +, +, +, - , - , - , +, - , +, +, - , - , +, +, +], [0, +, +, +, - , +, - , - , +, +, +, +, - , +, +, +], [- , +, - , - , - , +, +, - , +, +, +, - , - , - , +]
84	[0, - , - , +, +, - , - , +, - , - , +, - , - , +, - , -], [- , - , +, +, +, - , +, - , - , - , - , - , - , +, - , +, +, -], [0, +, - , +, - , - , +, +, - , - , - , - , +, +, - , +, - , +], [+ , - , +, +, +, - , +, +, - , - , - , +, +, - , +, +, +, -]

References

1. **M. J. E. Golay.** Complementary Series. *IRE Trans. Inform. Theory*, 1961, vol. IT-7, pp. 82–87.
2. **J. Seberry and M. Yamada.** Hadamard Matrices, Sequences, and Block Designs. In *Contemporary Design Theory: A Collection of Surveys*. J. H. Dinitz and D. R. Stinson (eds). John Wiley and Sons, 1992, pp. 431–560.
3. **W. de Launey and D. Flannery.** Algebraic Design Theory. *Mathematical Surveys and Monographs*, American Mathematical Society, Providence, R. I., 2011, vol. 175. 298 p.
4. **D. Ž. Đoković.** Note on Periodic Complementary Sets of Binary Sequences. *Designs, Codes and Cryptography*, 1998, no. 13, pp. 251–256.
5. **D. Ž. Đoković, I. S. Kotsireas.** Periodic Golay Pairs of Length 72. In *Algebraic Design Theory and Hadamard Matrices*, ADTHM, Lethbridge, Alberta, Canada, July 2014. C. J. Colbourn (ed). Springer, 2015, pp. 83–92.
6. **N. Ito.** On Hadamard Groups IV. *Journal of Algebra*, 2000, no. 234, pp. 651–663.
7. **N. Ito.** On Hadamard Groups III. *Kyushu J. Math.*, 1997, no. 51, pp. 369–379.
8. **B. Schmidt.** Williamson Matrices and a Conjecture of Ito's. *Designs, Codes and Cryptography*, 1999, no. 17, pp. 61–68.
9. **Balonin N. A., Đoković D. Ž.** Symmetry of Two Circulant Hadamard Matrices and Periodic Golay Pairs. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2015, no. 3(76), pp. 2–16. doi:10.15217/issn1684-8853.2015.3.2 (In Russian).
10. **H. J. Ryser.** Combinatorial Mathematics. *The Carus Mathematical Monographs*. Published by The Mathematical Association of America, New York, John Wiley and Sons, 1963, no. 14, p. 162.
11. **R. J. Turyn.** Hadamard Matrices, Baumert-Hall Units, four Symbol Sequences, Pulse Compression and Surface Wave Encodings. *J. Combin. Theory*, 1974, no. 16, pp. 313–333.
12. **P. Delsarte, J. M. Goethals, and J. J. Seidel.** Orthogonal Matrices with Zero Diagonal. II. *Can. J. Math.*, 1971, vol. XXIII, no. 5, pp. 816–832.
13. **K. T. Arasu, Y. Q. Chen, and A. Pott.** Hadamard and Conference Matrices. *Journal of Algebraic Combinatorics*, 2001, no. 14, pp. 103–117.
14. **B. Schmidt and M. M. Tan.** Construction of Relative Difference Sets and Hadamard Groups. *Designs, Codes and Cryptography*, 2014, no. 73, pp. 105–119.
15. **D. Ž. Đoković.** Equivalence Classes and Representatives of Golay Sequences. *Discrete Math.*, 1998, no. 189, pp. 79–93.
16. **R. E. A. C. Paley.** On Orthogonal Matrices. *J. Math. and Phys.*, 1933, no. 12, pp. 311–320.
17. **J. M. Goethals and J. J. Seidel.** Orthogonal Matrices with Zero Diagonal. *Can. J. Math.*, 1967, no. 19, pp. 1001–1010.
18. **W. H. Holzmann, H. Kharaghani, and B. Tayfeh-Rezaie.** Williamson Matrices up to Order 59. *Designs, Codes and Cryptography*, 2008, no. 46, pp. 343–352.
19. **R. G. Stanton and R. C. Mullin.** On the Nonexistence of a Class of Circulant Balanced Weighing Matrices. *SIAM J. Appl. Math.*, 1976, no. 30, pp. 98–102.

УДК 519.614

doi:10.15217/issn1684-8853.2015.5.2

Негапериодические пары Голея и матрицы АдамараН. А. Балонин^а, доктор техн. наук, профессор, korbendfs@mail.ruД. Ж. Джокович^б, доктор наук, профессор, djokovic@uwaterloo.ca^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, 67, Б. Морская ул., 190000, Санкт-Петербург, РФ^бУниверситет Ватерлоо, Ватерлоо, Онтарио, Канада

Цель: показать, что по аналогии с ординарными и периодическими парам Голея существуют и негапериодические пары Голея (впервые они появились под другим именем в трудах Н. Ито). **Методы:** исследуется конструкция адамаровых (и взвешенных) матриц, состоящая из двух негациклических блоков (2N-типа). Матрицы Адамара 2N-типа эквивалентны негапериодическим парам Голея. **Результаты:** показано, что, во-первых, если матрица Адамара имеет форму матрицы Теплица, то она должна быть либо циклической, либо негациклической. Во-вторых, произведение Тюринга пар Голея расширяемо до более общего произведения: с его помощью можно периодические пары Голея длины g умножить на негапериодические пары Голея длины v , получая негапериодические пары Голея длины gv . В-третьих, гипотеза Ито о матрицах Адамара эквивалентна гипотезе о существовании негапериодических пар Голея для всех возможных четных значений их длины. **Практическая значимость:** матрицы Адамара имеют непосредственное практическое значение для задач помехоустойчивого кодирования, сжатия и маскирования видеоинформации.

Ключевые слова — матрицы Адамара, циклические матрицы, негациклические матрицы, периодические пары Голея, негапериодические пары Голея.

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

Стоимость годовой подписки (6 номеров) для подписчиков России — 4800 рублей,
для подписчиков стран СНГ — 5400 рублей, включая НДС 18%, таможенные и почтовые расходы.

Подписку на печатную версию журнала можно оформить в любом отделении связи по каталогу:

«Роспечать»: № 48060 — годовой индекс, № 15385 — полугодовой индекс,

а также через посредство подписных агентств:

«Северо-Западное агентство „Прессинформ“»

Санкт-Петербург, тел.: (812) 335-97-51, 337-23-05,

эл. почта: press@crp.spb.ru, zayavka@crp.spb.ru,

сайт: <http://www.pinform.spb.ru>

«МК-Периодика» (РФ + 90 стран)

Москва, тел.: (495) 681-91-37, 681-87-47,

эл. почта: export@periodicals.ru, сайт: <http://www.periodicals.ru>

«Информнаука» (РФ + ближе и дальше зарубежье)

Москва, тел.: (495) 787-38-73, эл. почта: informnauka3@yandex.ru,

сайт: <http://www.informnauka.com>

«Деловая пресса»

Москва, тел.: (495) 962-11-11, эл. почта: podpiska@delpress.ru,

сайт: <http://delpress.ru/contacts.html>

«Коммерсант-Курьер»

Казань, тел.: (843) 291-09-99, 291-09-47, эл. почта: kazan@komcur.ru,

сайт: <http://www.komcur.ru/contacts/kazan/>

«Урал-Пресс» (филиалы в 40 городах РФ)

Сайт: <http://www.ural-press.ru>

«Идея» (Украина)

Сайт: <http://idea.com.ua>

«BTL» (Узбекистан)

Сайт: <http://btl.sk.uz/ru/cat17.html> и др.

На электронную версию нашего журнала (все выпуски, годовая подписка, один выпуск, одна статья)

вы можете подписаться на сайтах НЭБ: <http://elibrary.ru>;

РУКОНТ: <http://www.rucont.ru>; ИВИС: <http://www.ivis.ru/>

Полнотекстовые версии журнала за 2002–2014 гг.

в свободном доступе на сайте журнала (<http://www.i-us.ru>),

НЭБ (<http://www.elibrary.ru>)

и Киберленинки (<http://cyberleninka.ru/>

journal/n/informatsionno-upravlyayuschiesistemy).

Печатную версию архивных выпусков журнала за 2003–2014 гг. вы можете заказать в редакции по льготной цене.

МЕТОДЫ КОНВЕРТАЦИИ 2D-ИЗОБРАЖЕНИЙ И ВИДЕО В СТЕРЕОСКОПИЧЕСКИЙ ФОРМАТ

Н. Н. Красильников^а, доктор техн. наук, профессор

О. И. Красильникова^а, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Постановка проблемы: известные методы, обеспечивающие качественную конвертацию 2D-изображений и видео в стереоскопический формат, имеют общий недостаток — ручную реализацию ряда этапов и, следовательно, являются дорогостоящими. **Цель:** разработка полуавтоматических и потому более экономичных методов конвертации 2D-изображений и видео в стереоскопический формат. **Результаты:** предложенные шесть методов конвертации основаны, во-первых, на положениях выдвинутой ранее и экспериментально подтвержденной гипотезы, согласно которой ощущение глубины пространства при наблюдении 3D-сцен и стереоскопических изображений создается благодаря «конвергенционно-дивергенционным» движениям глаз, а во-вторых, на следующем из нее выводе, что при генерации стереоизображений важны лишь области карты глубины, примыкающие к резким световым границам. Первый метод заключается в аппроксимации поверхностей объектов, представленных на конвертируемом изображении, небольшим количеством плоскостей. Второй метод основан на синтезе карты глубины небольшим набором примитивных поверхностей. Особенностью третьего метода конвертации является синтез карты глубины из ахроматического компонента самого конвертируемого изображения. Четвертый метод конвертации отличается от предыдущего тем, что карта глубины формируется из контуров ахроматического компонента конвертируемого изображения. В пятом и шестом методах при формировании карты глубины использован фактор движения объектов. **Практическая значимость:** разработанные методы обладают большим быстродействием и требуют меньших затрат ручного труда, поэтому их можно эффективно применять на практике для конвертации как отдельных изображений, так и видео.

Ключевые слова — 3D-изображение, стереоскопическое изображение, конвертация, 2D-изображение, видео, 3D-формат.

Введение

В настоящее время значительное внимание уделяется конвертации 2D-изображений и фильмов в так называемый «3D»-формат. Мы заключаем в кавычки термин «3D», поскольку на самом деле речь идет о конвертации 2D-изображений и фильмов в стереоскопический формат, а использование термина «3D»-формат носит, скорее, рекламный характер. Повышенный интерес к конвертации обусловлен отсутствием достаточного количества качественного контента в данном формате. Несмотря на то что в Интернете имеется множество предложений по конвертации 2D-изображений и видео в «3D»-формат, а также бесплатных приложений конвертации и приложений конвертации, предоставляемых за весьма умеренную плату, описания самих методов конвертации практически отсутствуют. В Интернете имеются упоминания лишь о двух технологиях конвертации, которые считаются основными [1, 2].

Первая технология заключается в том, что для конвертируемого 2D-изображения сначала рисуется карта глубины, т. е. по существу для каждого объекта, представленного на этом изображении, рисуется силуэт, залитый серыми тонами разной светлоты. Затем, с использованием конвертируемого 2D-изображения и карты гл-

бины, рассчитываются левое и правое изображения стереопары.

При использовании второй технологии в 3D-среде 3D-графического редактора, например 3Ds Max, создается виртуальная модель сцены, повторяющая сцену, представленную на 2D-изображении. Затем на эту виртуальную модель проецируется 2D-изображение, и результат фотографируется двумя виртуальными камерами, смещенными по горизонтали на величину базы. Полученные таким образом изображения образуют стереопару.

Недостатком этих технологий является то, что их реализация требует больших затрат ручного труда. Действительно, вначале оператор вручную разделяет сцену конвертируемого изображения на составляющие ее объекты, т. е. выполняет сегментацию исходного 2D-изображения. Эта сегментация является семантической, т. е. смысловой, поскольку при этом определяется, какие пиксели изображения к какому объекту относятся. Последующие операции каждой из двух технологий также выполняются в ручном режиме. При использовании первой технологии после семантической сегментации области, занимаемые каждым из объектов на 2D-изображении, заполняются серым цветом, интенсивность которого (или его отдельных областей, если объект протяженный в глубь сцены) определяется оператором в соответствии с удаленностью этого объекта от

камеры. Применяя вторую технологию, после выполнения семантической сегментации виртуальную модель сцены, содержащей все объекты, оператор создает вручную.

Для автоматизации процесса конвертации необходимо решить две задачи, относящиеся к области искусственного интеллекта, которые до настоящего времени не решены. Это задача семантической сегментации изображений и задача автоматического определения расстояния от объекта до камеры на основании перспективных трансформаций изображений сегментированных объектов. В результате того, что описанные выше технологии базируются на ручном труде оператора, они оказались чрезвычайно дорогими. По словам режиссера Джеймса Кэмерона, создателя фильма «Титаник», стоимость конвертации этого фильма с использованием технологии рисования карт глубины составила около \$18 млн и потребовала работы 300 специалистов в течение 60 недель [2].

Стремление автоматизировать конвертацию 2D-изображений и видео в «3D»-формат привело к появлению ряда упрощенных методов, возможность разработки которых положили исследования, опубликованные в работах [3–6]. Некоторые из этих методов рассмотрены в настоящей статье.

Метод конвертации, основанный на аппроксимации 2D-изображения набором плоскостей

В основе метода, рассматриваемого здесь, лежат результаты наших исследований [9], по итогам которых было показано, что:

— ощущение глубины пространства создается благодаря тому, что при наблюдении 3D-сцен и стереоскопических изображений глаза зрителя при сканировании постоянно совершают «конвергенционно-дивергенционные» движения, одновременно фиксируясь вблизи резких световых границ составляющих их объектов;

— стереоскопические изображения допускают большие искажения параллакса без возникновения заметных зрителю искажений, из чего

следует допустимость больших погрешностей при создании карт глубины.

Отличием разработанного нами метода от метода, в котором применяется виртуальная модель сцены, представленной на 2D-изображении, является то, что при ее создании выполняется весьма грубая аппроксимация поверхностей составляющих ее объектов малым количеством плоскостей [7, 8]. Благодаря этому существенно упрощается процедура моделирования и в то же время возникающие погрешности параллакса для зрителя остаются незаметными в силу отмеченного выше свойства зрения. В качестве поясняющего примера на рис. 1, а в черно-белом варианте приведено 2D-изображение, на котором белыми линиями показаны границы аппроксимирующих плоскостей, которыми аппроксимируются стены, пол и потолок помещения, а на рис. 1, б — карта глубины, которую представляют эти пять плоскостей.

Светлые области карты глубины соответствуют частям сцены, близко расположенным к зрителю, а темные области — удаленным частям сцены.

Другим отличием данного метода является способ получения стереопары, которая рассчитывается на основе полученной карты глубины. Вначале вычисляются значения параллакса p по известной формуле

$$p = \frac{b(d_2 - d_1)}{d_2},$$

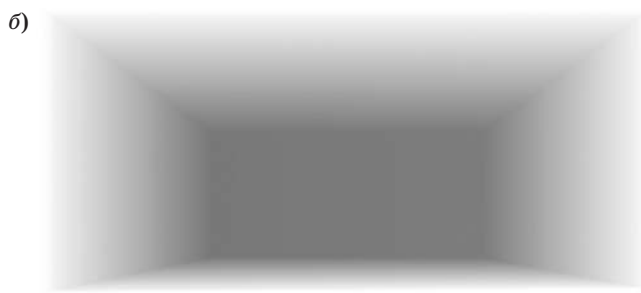
где b — база (расстояние между оптическими осями); d_1 — расстояние от зрителя до экрана, на котором наблюдается стереоскопическое изображение; d_2 — расстояние от зрителя до наблюдаемого виртуального объекта (берется из карты глубины). После этого рассчитываются компоненты изображений стереопары:

— левого

$$L_{LR}(x, y) = L_R(x - p/2, y);$$

$$L_{LG}(x, y) = L_G(x - p/2, y);$$

$$L_{LB}(x, y) = L_B(x - p/2, y);$$



■ Рис. 1. Исходное изображение, подлежащее конвертации (а), и его карта глубины (б)

— правого

$$L_{RR}(x, y) = L_R(x + p/2, y);$$

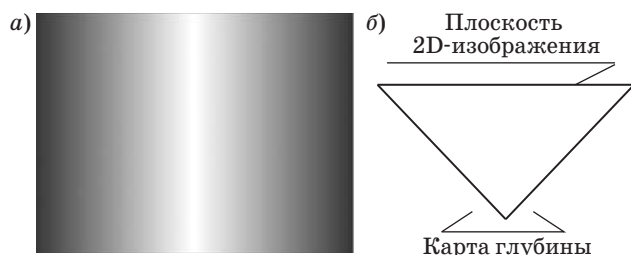
$$L_{RG}(x, y) = L_G(x + p/2, y);$$

$$L_{RB}(x, y) = L_B(x + p/2, y).$$

Здесь $L_{LR}(x, y)$, $L_{LG}(x, y)$, $L_{LB}(x, y)$ — зависимость интенсивности излучения от координат x и y для левого компонента в красном, зеленом и синем цветах; $L_{RR}(x, y)$, $L_{RG}(x, y)$, $L_{RB}(x, y)$ — то же самое для правого компонента. В приведенных формулах для простоты изложения изображения рассматривались как непрерывные функции координат x и y . Однако реальные изображения, представленные набором пикселей, дискретны в пространстве. Вследствие пространственных преобразований, обусловленных зависимостью параллакса от пространственных координат, в одних зонах стереоскопических изображений имеют место растяжения в горизонтальном направлении, а в других — сжатия. В результате этого на изображениях возникают места, в которых интенсивность пикселей не определена. Эта неопределенность устраняется методом интерполяции в горизонтальном направлении. В работе [10] на рис. П1 и П2 приведены исходное цветное 2D-изображение и результат его конвертации в «3D»-формат описанным методом. Следует особо отметить, что для изображений сцен, в которых отсутствуют объекты, заслоняющие собой какие-либо части изображенной сцены, полученная карта глубины достаточно точно отражает реальную ситуацию, что позволяет при расчете стереоскопического изображения изменять ракурс наблюдения в довольно широких пределах. Рис. П3 и П4 [10] демонстрируют это.

Метод конвертации 2D-изображения в «3D»-изображение, основанный на задании примитивной карты глубины

Следующий метод, предложенный нами, основан на задании в качестве карты глубины произвольной поверхности, в простейшем случае — двух плоскостей, ориентированных перпендикулярно плоскости горизонта (рис. 2, а, б).



■ Рис. 2. Примитивная карта глубины (а) в виде двух плоскостей, ориентированных перпендикулярно плоскости горизонта (б)

Карта глубины, заданная таким образом, обеспечивает лишь разные значения параллаксов для пикселей, имеющих различное горизонтальное расположение. Но благодаря тому, что при просмотре такого изображения имеют место «конвергенционно-дивергенционные» движения зрительных осей, будет возникать ощущение глубины пространства [9]. Естественно, вследствие того, что карта глубины задана произвольно, значения параллаксов стереоскопического изображения оказываются сильно искаженными, следовательно, расстояния до объектов, которые зрительная система «вычисляет» через значения параллаксов, могут существенно отличаться от расстояний, которые она вычисляет путем сопоставления масштабов объектов, расположенных на различных расстояниях, тем не менее в большинстве случаев эти расхождения остаются незамеченными зрителем. В этом легко убедиться, обратившись к рисункам сайта [11], на которых приведены исходные 2D-изображения и стереоскопические изображения, представляющие результат их конвертации описанным способом. В качестве карты глубины могут быть использованы также фрагменты цилиндрических, гауссовых и других поверхностей. В отличие от метода, описанного выше, при расчете стереоскопического изображения попытка изменить ракурс наблюдения даже в небольших пределах в данном случае приводит к недопустимо большим геометрическим искажениям.

Метод конвертации, основанный на использовании карты глубины, сформированной из исходного изображения

Недостатком метода, рассмотренного в предыдущем разделе, является то, что карта глубины формируется вне какой-либо связи с объектами сцены. В результате этого возникает упорядоченное изменение параллакса в направлении, которое определяется только положением задаваемой произвольной поверхности, используемой как карта глубины, и которое не связано с положением объектов в пространстве. Однако для ряда изображений удастся избежать этого недостатка, формируя карту глубины из самого конвертируемого изображения. Примером таких изображений являются изображения, на которых сильно выражен эффект воздушной перспективы. Используя то обстоятельство, что близко расположенные объекты имеют больший контраст, чем удаленные, и формируя из ахроматических компонентов таких изображений путем их небольшого размытия карты глубины, удастся получать на сформированных таким методом картах сильное различие по интенсивности между областями,

которые соответствуют близко расположенным объектам, и областями, которые соответствуют удаленным объектам, а следовательно, разнести их по дальности в правильном направлении. В результате многочисленных экспериментов было установлено, что наилучший результат достигается при размытии ахроматического компонента конвертируемого изображения в горизонтальном направлении путем его свертки с гауссовой функцией

$$L_d(n, k) = \sum_{n'=-\frac{N_n-1}{2}}^{\frac{N_n-1}{2}} L_{ca}(n+n', k)h(n'),$$

где $L_{ca}(n, k)$ — ахроматический компонент конвертируемого изображения; k — номер строки; n — номер столбца; N_n — протяженность импульсной функции;

$$h(n) = k_1 \exp(-n/n_0).$$

Здесь k_1 — нормирующий множитель; n_0 — параметр, определяющий ширину импульсной функции. В качестве примера, иллюстрирующего эффективность описанного метода, на рис. П5 [10] приведено изображение, которое затем было конвертировано в анаглифные стереоскопические изображения методом, основанным на задании примитивной карты глубины (рис. П6, [10]), и методом, в котором карта глубины формировалась из исходного изображения в соответствии с приведенным описанием (рис. П7, [10]). Сопоставление изображений, представленных на рис. П6 и П7 [10], показывает, что использование дополнительной информации, которая заключена в исходном изображении в виде различия контраста, позволило существенно усилить эффект объема стереоскопического изображения. Так же, как и в случае метода конвертации, основанного на задании примитивной карты глубины, при использовании данного метода попытка изменить ракурс наблюдения при расчете стереоскопического изображения приводит к недопустимо большим геометрическим искажениям.

Метод конвертации, основанный на использовании карты глубины, сформированной из контуров исходного изображения

При выполнении экспериментов, описанных в предыдущем разделе, было обращено внимание на то, что карта глубины в своих внутренних областях, соответствующих областям объекта с плавным изменением яркости и цвета и отстоящих от световых границ объекта на расстояние, превышающее значение величины погрешности фиксации зрительных осей относительно вертикальных световых границ объекта, может иметь произвольные значения интенсивности (яркости), и это не сказывается на формируемом стереоскопическом изображении. Кроме того, если карту глубины конкретного объекта расширить за его пределы, но так, чтобы она не вторглась в зону световых границ другого объекта, это также не скажется на формируемом стереоскопическом изображении. Эти факты не являются неожиданными и объясняются тем, что интенсивность (яркость) карты глубины важна лишь вблизи световых границ объекта, поскольку она в конечном итоге определяет, насколько должна быть смещена эта граница при расчете компонентов стереоскопического изображения вследствие параллакса. Это полностью согласуется с выдвинутой нами гипотезой [9], которая объясняет механизм создания у зрителя ощущения глубины наблюдаемой сцены. Данное свойство карты глубины может быть использовано на практике при конвертации 2D-изображений в стереоскопические изображения в процессе редактирования карты глубины. Получение контурного препарата из конвертируемого изображения — процедура не сложная, а редактировать карту глубины в виде контурного изображения значительно проще, чем редактировать карту глубины, представленную в виде полутонового изображения. Поясним сказанное на примере конвертации изображения, ахроматическая версия которого приведена на рис. 3, а, с дополнительным требованием сильного разделения переднего и заднего планов.



■ Рис. 3. Исходное изображение (а) и карта глубины (б), сформированная из контуров исходного изображения

То есть при конвертации требуется усилить разделение по глубине сцены переднего и заднего планов.

Поставленная задача рассматриваемым методом решается следующим образом. Вначале из исходного изображения получается контурный препарат одним из известных методов. Лучше всего для этой цели использовать метод Канни, поскольку при этом на контурном препарате получается минимальное количество разрывов и отдельных штрихов и точек. После этого в данном конкретном случае следует устранить контуры, созданные объектами заднего плана. Работая с контурами, это сделать легко. Затем отредактированный таким образом препарат нужно подвергнуть гауссовой фильтрации в горизонтальном и вертикальном направлениях с целью устранить остаточные разрывы контуров и расширения контурных линий, а затем после двустороннего ограничения присвоить ему соответствующий уровень яркости. Полученный результат можно использовать как карту глубины (рис. 3, б). На рис. П8 и П9 [10] приведены исходное изображение и результат его конвертации в анаглифное стереоскопическое изображение. Из рис. П9 [10] видно, что полученный результат впечатляет.

Метод конвертации, основанный на использовании фактора движения при формировании карты глубины

В ряде случаев дополнительным признаком, отличающим задний план сцены от переднего плана, на котором сосредоточены объекты, является движение объектов. Это может быть использовано при формировании карты глубины. Переходя к рассмотрению предлагаемого метода конвертации, основанного на использовании данного признака, будем считать видеокамеру, посредством которой производится съемка, неподвижной. В качестве исходного препарата, из которого синтезируется карта глубины в этом методе, используются абсолютные значения разности интенсивности (яркости) ахроматических компонентов конвертируемого кадра и кадра, который ему предшествует:

$$L_{\Pi}(k, n, i) = |L_c(k, n, i) - L_c(k, n, i-1)|,$$

где $L_{\Pi}(k, n, i)$ — исходный препарат; $L_c(k, n, i)$ — распределение яркости в конвертируемом кадре изображения; $L_c(k, n, i-1)$ — распределение яркости в кадре, предшествующем конвертируемому кадру; k — номер строки; n — номер столбца (пикселя в строке); i — номер кадра. Исходный препарат не может быть непосредственно использован в качестве карты глубины по двум причинам. Первой причиной является наличие флуктуационного шума, а также небольшого разност-

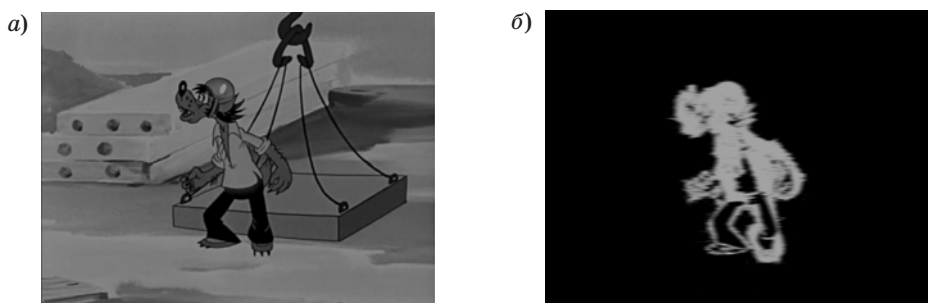
ного сигнала, обусловленного небольшими движениями видеокамеры, особенно при съемке с рук, которые в конечном итоге будут создавать ложные, изменяющиеся во времени приращения параллакса. Для того чтобы избавиться от этого эффекта, исходный препарат $L_{\Pi}(k, n, i)$ необходимо подвергнуть ограничению снизу, в результате чего получаем первый промежуточный препарат $L_{\Pi1}(k, n, i)$:

$$L_{\Pi1}(k, n, i) = \begin{cases} L_{\Pi}(k, n, i) & \text{при } L_{\Pi}(k, n, i) \geq P_{\Pi} \\ 0 & \text{при } L_{\Pi}(k, n, i) < P_{\Pi} \end{cases}, \quad (1)$$

где P_{Π} — нижний порог. Вторая причина, по которой, даже избавившись от флуктуационного шума и разностного сигнала, обусловленного небольшими движениями видеокамеры, нельзя использовать препарат $L_{\Pi1}(k, n, i)$ в качестве карты глубины, является сильная зависимость интенсивности $L_{\Pi1}(k, n, i)$ не только от яркости, но и от скорости движения объекта и его отдельных частей, которая непрерывно изменяется внутри одной сцены. Примером может служить фигура человека, который, будучи сам неподвижным, лишь двигает рукой. В этом случае, рассматривая препарат $L_{\Pi1}(k, n, i)$ как карту глубины, обнаружим на нем только нечто вроде силуэта руки, так как лишь она движется, а все остальные части изображения неподвижны. А поскольку от скорости движения руки зависит интенсивность этой части препарата (яркость), то величина рассчитанного по нему параллакса также будет определяться этой скоростью. В результате кажущаяся дальность руки будет изменяться вместе с изменением скорости ее движения, при этом она может на стереоскопическом изображении сильно отдаляться от своего владельца. Выход из этого положения был найден в сильном ограничении интенсивности препарата сверху путем его сравнения с порогом, который устанавливается на малом уровне. Таким образом, следующая вычислительная операция при формировании карты глубины, которая делает независимой ее интенсивность от скорости движения объекта, заключается в ограничении первого промежуточного препарата сверху:

$$L_{\Pi2}(k, n, i) = \begin{cases} L_{\Pi1}(k, n, i) & \text{при } L_{\Pi1}(k, n, i) \leq P_{\Pi} \\ P_{\Pi} & \text{при } L_{\Pi1}(k, n, i) > P_{\Pi} \end{cases}, \quad (2)$$

где $L_{\Pi2}(k, n, i)$ — второй промежуточный препарат; P_{Π} — верхний порог. И заключительной вычислительной операцией формирования карты глубины является размытие по горизонтали и вертикали второго промежуточного препарата подобно тому, как это сделано в методе кон-



■ Рис. 4. Исходное изображение (а) и карта глубины (б) в методе конвертации, основанном на использовании фактора движения при формировании карты глубины

вертации, основанном на использовании карты глубины, сформированной из исходного изображения, и присвоение ему соответствующего уровня яркости. При этом значение интенсивности (яркости) карты глубины, которая в конечном итоге определяет расположение объекта по дальности, устанавливается оператором в диалоговом режиме. На рис. 4, а в полутоновом черно-белом формате приведено исходное изображение, а на рис. 4, б — соответствующая ему карта глубины, полученная описанным выше способом. В приложении работы [10] на рис. П10 показан цветной вариант изображения, приведенного на рис. 4, а на рис. П11 — его стереоскопическое изображение в анаглифном формате.

Формирование карты глубины в случае движения камеры

Продолжая рассмотрение метода конвертации, основанного на использовании признака движения объекта, будем считать, что видеокамера, посредством которой производится съемка, также движется, но с другой скоростью, нежели объект. В этом случае для того, чтобы выделить на карте глубины только движущийся объект, необходимо скомпенсировать смещение от кадра к кадру изображения снимаемой сцены за исключением объекта, который мы хотим выделить. Другими словами, необходимо в изображениях двух следующих друг за другом кадров фоновую часть сделать неподвижной (разумеется, за исключением небольших периферийных областей, которые в соседних кадрах не совпадают вследствие движения камеры). При горизонтальном движении камеры (наиболее часто встречающийся случай) в изображениях кадров, следующих друг за другом, необходимо выделить по одной строке с одинаковым номером k_0 , которые не пересекают выделяемый движущийся объект (или объекты, если их несколько), и путем сравнения их содержимого определить взаимное смещение фоновой части изображений в этом направлении. Аналогичным образом определяется величина

взаимного смещения и при движении камеры в любом другом направлении. Все вычислительные операции целесообразно выполнять с ахроматическими компонентами изображений. Для конкретности дальнейшего изложения будем считать, что видеокамера при съемке смещается в горизонтальном направлении.

Вначале вычисляются суммы абсолютных значений разностей яркостей пикселей в строках k_0 сравниваемых изображений $L_c(k_0, n, i)$ и $L_c(k_0, n, i - 1)$ при различных величинах их взаимного смещения на r пикселей в направлении движения камеры, которые записываются в массив $M(r)$:

$$M(r) = \sum_{n=r_m}^{N-r_m} |L_c(k_0, n, i) - L_c(k_0, n + r, i - 1)|,$$

где N — число пикселей (столбцов) в изображении; k_0 — номер строки в сравниваемых изображениях; r_m — максимальный ожидаемый интервал взаимного смещения фоновой части изображений в ту или иную сторону в пикселях. При этом $-r_m \leq r \leq r_m$. Затем путем перебора элементов массива $M(r)$ находится элемент массива с минимальным значением, индекс которого r_0 представляет собой величину взаимного смещения изображений текущего и предыдущего кадров, обусловленных движением снимающей видеокамеры. После этого, как и в предыдущем разделе, по формуле

$$L_{\pi}(k, n, i) = |L_c(k, n, i) - L_c(k, n + r_0, i - 1)|$$

определяется исходный препарат $L_{\pi}(k, n, i)$ для области изображения, занимаемой выделяемым движущимся объектом, из которого затем по формулам (1) и (2) синтезируются первый $L_{\pi 1}(k, n, i)$ и второй $L_{\pi 2}(k, n, i)$ промежуточные препараты, как это описано в разделе «Метод конвертации, основанный на использовании фактора движения при формировании карты глубины». Заключительным этапом формирования карты глубины является размытие по горизонтали и вертикали второго промежуточного препарата



■ Рис. 5. Текущий (а) и предшествующий (б) кадры и карта глубины (в) в методе конвертации, основанном на использовании фактора движения при формировании карты глубины в случае движения камеры

и присвоения соответствующего уровня яркости. Значение интенсивности (яркости) карты глубины, которая в конечном итоге определяет расположение объекта по дальности, устанавливается оператором в диалоговом режиме. На рис. 5, а в полутоновом черно-белом формате приведено исходное изображение текущего кадра, на рис. 5, б — предшествующего ему кадра, а на рис. 5, в — соответствующая ему карта глубины, полученная описанным выше способом. В приложении [10] на рис. П12 показан цветной вариант изображения, приведенного на рис. 5, а, а на рис. П13 — его стереоскопическое изображение в анаглифном формате.

Заключение

Последнее время характеризуется повышенным интересом к так называемым 3D-кино и телевидению, но отсутствие качественного контента сдерживает развитие этих областей медиаиндустрии. Вместе с тем съемка фильмов в этом формате по сравнению с обычными, в формате 2D, требует значительных финансовых затрат. Не многим дешевле обходится конвертация уже снятых 2D-фильмов в 3D-формат, поскольку существующие технологии требуют больших за-

трат ручного труда. Выходом из сложившейся ситуации является разработка методов конвертации, базирующихся на различных особенностях конкретного конвертируемого контента: факторе движения; различии в резкости световых границ объектов, составляющих сцену; наличии в сцене объектов, поверхности которых могут быть аппроксимированы плоскостями; наличии характерных контуров; явно выраженной перспективе в изображении конвертируемой сцены и т. д. Имея пакет компьютерных приложений, в основу которых положен набор таких методов, роль оператора можно свести только к выбору надлежащего приложения для конвертации конкретного клипа (фрагмента видео, содержащего определенный тип 2D-изображений) и подбору необходимых параметров конвертации. Опыт работы с приложениями, основанными на описанных методах, показал реальность подобного решения. Отметим также, что в наших экспериментах с описанными в статье приложениями время конвертации одного кадра видео размером 1920×1080 не превышало 1 с для любого из них. На сайте [12] в качестве примера приведено видео, при конвертации которого был применен метод, основанный на использовании карты глубины, сформированной из исходного изображения.

Литература

1. Конвертация 2d в 3d. Что такое псевдостерео? <http://www.televizor-3d.ru/konvertaciya-2-d-v-3d.html> (дата обращения: 10.10.2014).
2. Почему «Титаник» стоит смотреть в стерео 3D. <http://total3d.ru/diy/92707/> (дата обращения: 10.10.2014).
3. Красильников Н. Н. Метод получения 3D-изображений, основанный на диффузном отражении света сканируемыми объектами // Информационно-управляющие системы. 2009. № 6(43). С. 7–11.
4. Красильников Н. Н., Красильникова О. И. Получение трехмерного изображения объекта путем измерения интенсивности диффузного отражения света различными точками его поверхности // Оптический журнал. 2010. Т. 77. № 6. С. 19–24.
5. Красильников Н. Н., Красильникова О. И. Определение координаты глубины по 2D изображению // Оптический журнал. 2011. Т. 78. № 12. С. 30–33.
6. Красильников Н. Н., Красильникова О. И. Исследование погрешностей определения координаты глубины при 3D-сканировании методом, основанным на диффузном отражении света // Информационно-управляющие системы. 2012. № 3(58). С. 2–8.
7. Красильников Н. Н. Метод формирования 3D-изображения сцены по одной фотографии // Информационно-управляющие системы. 2013. № 3. С. 2–7.
8. Красильников Н. Н., Красильникова О. И. Метод конвертации 2D-изображения в стереоскопическое

3D-изображение // Оптический журнал. 2014. Т. 81. № 2. С. 20–28.

9. Красильников Н. Н. Механизмы и погрешности восприятия 3D-пространства зрительной системой // Информационно-управляющие системы. 2014. № 6(73). С. 13–19.
10. Приложение к статье «Методы конвертации 2D изображений и видео в стереоскопический формат». http://cv.guap.ru/krasilnikov_nn/files/pril006 (дата обращения: 10.07.2014).

11. Приложение к исследованиям нового метода конвертации 2D-изображений в 3D-изображения. http://cv.guap.ru/krasilnikov_nn/files/pril003 (дата обращения: 01.07.2015).

12. Приложение к исследованиям нового метода конвертации 2D-изображений в 3D-изображения. http://cv.guap.ru/krasilnikov_nn/files/pril004 (дата обращения: 01.07.2015).

UDC 612.843.721

doi:10.15217/issn1684-8853.2015.5.18

Methods of Converting 2D Images and Videos into Stereoscopic Format

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

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

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

Purpose: The known methods of providing high-quality conversion of 2D images and videos into stereoscopic format have a common drawback: some stages require manual operations which make this conversion costly. The aim of this work is developing more economical semi-automatic conversion methods. **Results:** The six proposed conversion methods are based on the hypothesis (put forward by us and experimentally confirmed) that space depth perception when observing 3D scenes and stereoscopic images is provided by "converging-diverging" eye movements. This hypothesis has a corollary that when generating stereo images, only the fields of the depth map adjacent to the sharp light borders are important. The first method is approximating the surfaces of the objects represented on the converted image by a small number of planes. The second method is synthesizing the depth map by a small set of primitive surfaces. The main feature of the third conversion method is synthesizing the depth map from the achromatic component of the converted image. The fourth method differs from the previous one by generating the depth map out of the contours of the achromatic component of the converted image. The fifth and sixth methods use the objects movement factor to form the depth map. **Practical relevance:** The proposed methods, as compared to the conventional ones, have greater performance and require less manual labor. They can be effectively used in practice, to convert either single images or videos.

Keywords — 3D Image, Stereoscopic Image, Conversion, 2D Image, Video, 3D Format.

References

1. *Konvertatsiya 2d v 3d. Chto takoe pseudostereo?* [Converting 2D to 3D. What is Pseudostereo?]. Available at: <http://www.televizor-3d.ru/konvertatsiya-2-d-v-3d.html> (accessed 10 October 2014).
2. *Pochemu «Titanik» stoit smotret' v stereo 3D* [Why «Titanic» is Worth Watching in Stereo 3D]. Available at: <http://total3d.ru/diy/92707/> (accessed 10 October 2014).
3. Krasilnikov N. N. A Method of 3D Scanning Based on Scanned Objects' Light Radiosity. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2009, no. 6(43), pp. 7–11 (In Russian).
4. Krasilnikov N. N., Krasilnikova O. I. Obtaining a Three Dimensional Image of an Object by Measuring the Intensity of the Diffuse Reflection of Light by Various Points of its Surface. *Opticheskii zhurnal*, 2010, vol. 77, no. 6, pp. 19–24 (In Russian).
5. Krasilnikov N. N., Krasilnikova O. I. Determining the Depth Coordinate from a 2D Image. *Opticheskii zhurnal*, 2011, vol. 78, no. 12, pp. 30–33 (In Russian).
6. Krasilnikov N. N., Krasilnikova O. I. The Investigation of Measurement Errors of Depth Coordinate during 3D Scanning by the Method Based on Diffuse Reflection of Light. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 3(58), pp. 2–8 (In Russian).
7. Krasilnikov N. N. The Method of Generating a 3D-Image Scene Based on a Single Photograph. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 3(64), pp. 2–7 (In Russian).
8. Krasilnikov N. N., Krasilnikova O. I. The Method of Converting a 2D Image into a Stereoscopic 3D Image. *Opticheskii zhurnal*, 2014, vol. 81, no. 2, pp. 20–28 (In Russian).
9. Krasilnikov N. N. Mechanisms and Inaccuracies of 3D Spatial Perception by Visual System. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2014, no. 6(73), pp. 13–19 (In Russian).
10. *Prilozhenie k stat'e "Metody konvertatsii 2D izobrazhenii i video v stereoskopicheskiy format"* [The Appendix to the Article "Methods of Converting a 2D Image and Video into a Stereoscopic Format"]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril006 (accessed 10 April 2014).
11. *Prilozhenie k issledovaniyam novogo metoda konvertatsii 2D-izobrazhenii v 3D-izobrazheniia* [Appendix to the Research of a New Method of Converting 2D Images into 3D Images]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril003 (accessed 01 July 2015).
12. *Prilozhenie k issledovaniyam novogo metoda konvertatsii 2D-izobrazhenii v 3D-izobrazheniia* [Appendix to the Research of a New Method of Converting 2D Images into 3D Images]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril004 (accessed 01 July 2015).

ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ БИОМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ ПОЛЬЗОВАТЕЛЕЙ НА ОСНОВЕ ДАННЫХ АЙТРЕКЕРА TOBII X2-30

А. В. Лямин^а, канд. техн. наук, доцент

Е. Н. Череповская^а, старший лаборант

^аСанкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, РФ

Цель: биометрическая идентификация на основе анализа движения глаз обладает высокой степенью защиты от подлога. Точность этого метода идентификации зависит от анализируемых параметров движения глаз, применяемых стимулов, а также частоты прибора регистрации траектории взгляда. Цель исследования заключается в определении параметров движения глаз, имеющих наиболее существенное значение для идентификации, а также стимулов, которые нужно использовать, чтобы выявить характерные особенности движения глаз конкретного человека на основе данных низкочастотного прибора регистрации траектории взгляда Tobii X2-30. **Результаты:** проведены экспериментальные исследования идентификации личности на основе анализа движения глаз, зафиксированного с помощью айтрекера Tobii X2-30, при воздействии двух различных стимулов. Данные, полученные с помощью систем регистрации траектории взгляда, в том числе и низкочастотных, могут быть использованы для идентификации личности человека. Анализ данных двух экспериментов показал, что ошибка идентификации может варьироваться от 0 до 40 % в зависимости от рассматриваемого параметра движения. В ходе обработки результатов второго эксперимента было выявлено, что анализ саккад обеспечивает более точную идентификацию по сравнению с анализом параметров, рассчитанных для фиксаций, за счет снятия ограничения на скорости и амплитуды движений глаза. По результатам второго эксперимента определено, что значение ускорения взгляда при рассмотрении саккад может быть использовано для идентификации личности. **Практическая значимость:** полученные в ходе проведенного экспериментального исследования результаты подтверждают возможность использования низкочастотных систем регистрации траектории взгляда в целях идентификации личности человека.

Ключевые слова — информационные системы, биометрические методы идентификации, анализ движения глаз.

Введение

Разработка надежных, недорогих и удобных методов идентификации пользователей — серьезная проблема из-за растущего спроса на услуги, которые предоставляются информационными системами. Под идентификацией в информационных системах понимают установление личности пользователя (идентификатора) по характерным признакам (факторам). Традиционные методы идентификации делятся на два типа. Первый тип предполагает наличие у пользователя специального компактного устройства: USB-ключа, пластиковой карты и т. п. Однако эти устройства легко украсть, их можно потерять или повредить. Второй тип методов основывается на специальных знаниях пользователя, когда система запрашивает ввод пароля или же воспроизведение некоторых графических комбинаций. Несмотря на то что методы данного типа широко используются в информационных системах, они не лишены недостатков. Во-первых, они не предоставляют достоверной информации о пользователе, так как невозможно определить личность человека, основываясь лишь на проверке правильности введенного пароля. Во-вторых, они требуют от пользователей запоминания большого количества лишней информации, например паролей, вопросов и от-

ветов для их восстановления, графических комбинаций.

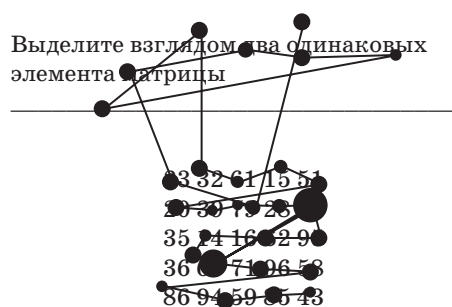
Отмеченные недостатки традиционных методов идентификации способствуют увеличению интереса к методам идентификации на основе биометрических данных. По сравнению с традиционными методами идентификации биометрические методы обеспечивают более защищенные способы контроля доступа на основе анализа физиологических особенностей (отпечатка пальца [1], формы ладони [2], геометрии лица [3], формы ушной раковины [4], узора радужной оболочки глаза [5], рисунка кровеносных сосудов на сетчатке глаза [6] и т. д.); поведенческих характеристик (походки [7], подписи [8], электронного почерка [9], голоса [10], траектории взгляда [11], электрической активности сердца [12] или мозга [13] и т. д.), а также какого-либо сочетания вышеперечисленных факторов (многофакторные методы идентификации).

Среди методов биометрической идентификации наиболее распространен метод по отпечаткам пальцев, который был успешно применен в многочисленных приложениях. Несмотря на то что все люди, как известно, имеют уникальные и неизменные отпечатки пальцев, метод идентификации по отпечаткам пальцев не защищен от подлога, поскольку человек может предна-

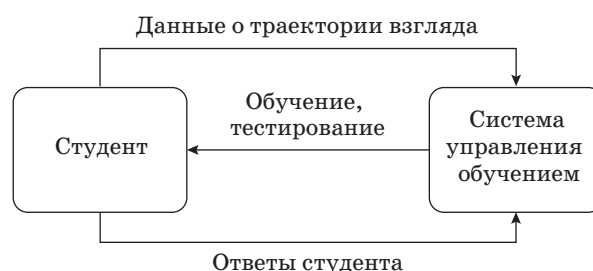
меренно оставить отпечатки своих пальцев где-нибудь, где их возможно будет скопировать. Для повышения уровня защиты от подлога в системах контроля доступа начали применять многофакторные методы биометрической идентификации, которые предполагают анализ одновременно нескольких факторов, например отпечатка пальца и узора радужной оболочки глаза [14]. Желание повысить уровень защиты от подлога инициировало также разработку методов биометрической идентификации, построенных на анализе поведенческих характеристик, которые невозможно или чрезвычайно сложно скопировать. К таким поведенческим характеристикам относят движения глаз.

Движения глаз содержат трудновоспроизводимую биометрическую информацию, которая невидима постороннему. Взгляд человека либо фиксируется на некоторой области в поле зрения (фиксация), либо быстро перемещается между фиксациями (саккада). Продолжительность фиксации составляет от 200 до 350 мс, продолжительность саккады не превышает 200 мс. На рис. 1 показана траектория взгляда человека, который осуществляет поиск двух одинаковых элементов в матрице размером 5×5. Видно, что траектория состоит из последовательности сменяющих друг друга фиксаций и саккад. Обычно для регистрации траектории взгляда пользователя применяют специальное устройство — айтрекер. Это устройство способно определять фиксации и саккады, координаты траектории взгляда, положение глаз в пространстве.

Известные методы биометрической идентификации на основе анализа движения глаз имеют низкую точность, что препятствует их широкому применению на практике [15]. Глаза человека представляют собой богатый источник информации, только частично используемый для биометрии. Например, неизвестно, какие параметры движения глаз имеют существенное значение для идентификации, какие стимулы позволяют наиболее эффективным способом выявить характерные свойства движения глаз человека.



■ Рис. 1. Визуализация фиксаций (точки) и саккад (линии) взгляда



■ Рис. 2. Схема взаимодействия студента с системой электронного обучения

Необходимы дополнительные исследования, чтобы понять, какую информацию могут предоставить глаза и как эта информация должна быть проанализирована.

Ряд исследований показал, что анализ траектории взгляда можно применять не только для идентификации личности пользователя, но и для оценки его функционального состояния и уровня навыков в различных областях деятельности [16, 17]. Например, методы анализа движения глаз могут широко использоваться в интеллектуальных системах электронного обучения (рис. 2, [18]), что позволит: снизить стоимость интернет-экзаменов, повысить доверие педагогического сообщества к результатам электронного обучения за счет непрерывной идентификации личности студента; повысить точность педагогических оценок и эффективность электронного обучения за счет непрерывного контроля функционального состояния студента [19, 20] и уровня его подготовки. Методы анализа движения глаз могут применяться при проведении интернет-экзаменов в компьютерных классах [21] и при дистанционном надзоре, так как уже сейчас многие персональные компьютеры оборудованы 3D-камерами, которые можно использовать в качестве айтрекеров.

В данной статье представлены результаты экспериментальных исследований биометрической идентификации пользователей на основе анализа движения глаз, зафиксированного с помощью айтрекера Tobii X2-30, при воздействии двух различных стимулов.

Этапы проведения исследования и метод анализа итоговых результатов

Участники и схема исследования

В экспериментальном исследовании были использованы два теста с различными стимулами. Испытуемые должны были выполнять поставленные перед ними задачи по обоим тестам. Данные были собраны с помощью айтрекера Tobii X2-30, частота дискретизации которого составляет 30 Гц. Каждая запись включала временную метку в миллисекундах, информацию о виде движения

■ Таблица 1. Схемы экспериментов

Идентификатор испытуемого		Количество повторов теста	
Первый эксперимент	Второй эксперимент	Первый эксперимент	Второй эксперимент
1	1	1	2
2	2	2	2
3	3	2	1
4	–	3	–
5	–	1	–

(саккада или фиксация), координаты обоих глаз, координаты взгляда и диаметры зрачков левого и правого глаза.

Цель первого эксперимента заключалась в анализе параметров, рассчитанных для фиксаций. Было подготовлено простое стимул-изображение для получения наиболее точных измерений. Испытуемым было необходимо в течение 20 с смотреть в центр показанного им изображения, где находился небольшой круг — точка фиксации. Полученные в ходе эксперимента данные были очищены от записей, расположенных на границе саккад и фиксаций. В эксперименте приняли участие пять испытуемых, не имеющих проблем со зрением, для которых было собрано в сумме девять записей. Схема прохождения эксперимента представлена в табл. 1.

Второй эксперимент включал в себя анализ как фиксаций, так и саккад. Стимул для данного эксперимента был выбран таким образом, чтобы предоставить точные данные для обоих видов движения. Для этой цели было выбрано изображение-реализация теста переплетенных линий А. Рея. Испытуемым было необходимо поочередно проследить каждую из представленных на рисунке линий и назвать цифры, связанные с левым и правым концом линии. В эксперименте приняли участие трое испытуемых, не имеющих проблем со зрением, для которых было собрано в сумме пять записей (см. табл. 1).

Анализ полученных данных

На основе собранных данных были рассчитаны необходимые для последующего анализа динамические параметры: линейная и угловая скорость, ускорение взгляда; скорость и ускорение изменения диаметра зрачков. Параметры были рассчитаны отдельно для левого и правого глаза для проведения сравнительного анализа.

Был использован тест парных выборок Колмогорова — Смирнова с пятипроцентным уровнем значимости для каждой пары записей. В качестве результата были получены таблицы вероятностей, содержащих информацию о степени схожести данных в каждой паре выборок для рас-

■ Таблица 2. Таблица вероятностей для скорости взгляда левого глаза для данных фиксаций, полученных в ходе второго эксперимента

	P_1	P_2	P_3	P_4	P_5
P_1	1	0,03	$2,3 \times 10^{-4}$	$9,37 \times 10^{-27}$	$3,22 \times 10^{-46}$
P_2		1	0,24	$1,31 \times 10^{-22}$	$1,2 \times 10^{-43}$
P_3			1	$2,63 \times 10^{-12}$	$6,67 \times 10^{-30}$
P_4				1	$1,03 \times 10^{-6}$
P_5					1

считанных параметров. Пример одной из полученных таблиц вероятностей приведен в табл. 2.

После этого полученные таблицы вероятностей и результатов теста Колмогорова — Смирнова для проверки нулевой гипотезы были проанализированы. Для каждого из двух экспериментов результаты были сопоставлены с ключами тестов — идентификационными кодами испытуемых и кодами, сгенерированными для каждого из них при обработке, что позволило определить погрешность идентификации.

Результаты исследования

Главная цель экспериментального исследования — определение того, какой вид движения глаз и какие вышеперечисленные динамические параметры движения левого и правого глаза позволяют получить наименьшую ошибку идентификации личности человека.

Первый эксперимент

Целью первого эксперимента был анализ записей для фиксаций. Для последующей обработки были сформированы три набора данных. Первый включал стандартные данные, полученные при помощи айтрекера, без проведения дополнительной обработки. Второй набор включал записи, нормированные по формуле

$$X' = \frac{X - \mu}{\sigma},$$

где X' — новое значение; X — предыдущее значение; μ — среднее значение и σ — стандартное отклонение выборки. Третий набор был сформирован данными между $1/4$ и $3/4$ квантилями с целью уменьшить возможный шум. Описанные типы наборов данных были проанализированы отдельно для всех фиксаций и для самой длинной фиксации.

Первичный анализ сформированных наборов данных показал, что наиболее значимые результаты были получены для первого типа обеих гипотез. Второй и третий наборы, записи в которых были статистически обработаны, не предоставили необходимую точность результатов, так как

все значения вероятностей были расположены близко к нулю, что автоматически означало отвержение гипотезы о принадлежности выборок к единой совокупности.

Наиболее точные результаты для первого эксперимента были получены при рассмотрении параметра угловой скорости взгляда. Значения ошибок первого и второго рода, полученные для этих данных, представлены в табл. 3.

Таблицы вероятностей для объединенных фиксаций и самой продолжительной фиксации были сравнимы по положению значимых вероятностей, однако значения различались. Результаты обработки данных для самой продолжительной фиксации представляли более высокие значения, что привело к увеличению погрешности. Далее рассматривались только результаты, полученные в ходе обработки записей объединен-

■ **Таблица 3.** Значения ошибок первого и второго рода, %, полученные в ходе анализа фиксаций взгляда при обработке данных первого эксперимента

Параметр	α -ошибка		β -ошибка	
	Левый глаз	Правый глаз	Левый глаз	Правый глаз
Угловая скорость	0		32	
Диаметр зрачка	80	100	0	0
Скорость зрачка	100	80	6	3
Скорость взгляда	80	80	13	6
Ускорение взгляда	80	80	16	16
Ускорение зрачка	80	100	3	3

■ **Таблица 4.** Результаты для параметра угловой скорости, рассчитанные в ходе обработки данных первого эксперимента

Первая запись	Вторая запись	Вероятность совпадения
P_1	P_2	0,69
P_1	P_3	0,19
P_1	P_6	0,56
P_1	P_9	0,81
P_2	P_6	0,44
P_2	P_7	0,27
P_2	P_9	0,43
P_3	P_6	0,81
P_3	P_9	0,37
P_4	P_7	0,29
P_4	P_8	0,34
P_6	P_9	0,8
P_7	P_8	0,67

ных фиксаций. Всего было зарегистрировано девять записей, закодированных как $P_1, P_2, \dots, P_9$. Результаты представлены в табл. 4.

Данная таблица предоставляет информацию о тех парах записей, для которых была установлена степень их совпадения. Анализируя ее, можно выделить следующие кластеры:

$$\begin{aligned} P_4 &= P_7 = P_8; \\ P_5; \\ P_3 &= P_6 = P_9 = P_1; \\ P_1 &= P_2 = P_9. \end{aligned}$$

Кластеры были сформированы на основе анализа связей между парами выборок. В первый кластер включены 4-, 7- и 8-я записи. Во второй кластер была включена только 5-я запись, так как связи между ней и остальными выборками не были обнаружены. Следовательно, 5-я запись была отделена и обозначена уникальной. Третий кластер состоит из четырех записей, а четвертый кластер — из трех.

После формирования кластеров было необходимо определить точность предоставляемых результатов для каждого из кластеров на основе ключей эксперимента:

- в первом кластере все представленные зависимости соответствуют реальным;
- запись во втором кластере действительно была уникальной;
- в третьем кластере только одна из зависимостей соответствует действительности;
- четвертый кластер содержит одну верную и одну неверную зависимость.

Погрешность, рассчитанная для представленных кластеров, составляет 38 %.

Следующий шаг обработки заключался в возможном уменьшении погрешности с выделением из рассматриваемой таблицы пар записей с вероятностью более 40 %. На основе этого были сформированы новые кластеры:

$$\begin{aligned} P_7 &= P_8; \\ P_5; \\ P_3 &= P_6 = P_9; \\ P_1 &= P_2 = P_9. \end{aligned}$$

Анализ их в сравнении с первым набором выявил, что:

- в первом кластере пропала одна верная связь;
- второй кластер остался неизменным;
- в третьем кластере пропала одна неверная зависимость;
- четвертый кластер остался неизменным.

Из сравнения новых кластеров с ключами эксперимента следует, что погрешность составила 33 %.

Анализ данных, полученных в ходе проведения первого эксперимента, показал, что наиболее

удачным параметром для осуществления идентификации личности по фиксациям взгляда является угловая скорость, погрешность использования которого может достигать 38 %.

Второй эксперимент

Целью второго эксперимента был анализ данных для двух видов движения глаз — саккад и фиксаций, однако главными рассматриваемыми параметрами являлись параметры саккад.

Было сформировано три группы данных. Первая из них включала в себя данные для фиксаций, вторая — данные для саккад, а третья группа объединяла данные для саккад и фиксаций.

Как и при обработке данных первого эксперимента, для фиксаций были реализованы два варианта обработки: объединенные данные всех фиксаций и данные для самой продолжительной фиксации, — в то время как для саккад был осуществлен только первый вариант обработки экспериментальных данных. Для объединенных данных были исключены точки, находящиеся на границе саккад и фиксаций.

Во время обработки результатов второго эксперимента были рассмотрены изначальные данные, без проведения предварительной статистической обработки. Были рассчитаны 11 параметров для каждой из пяти полученных записей.

Анализ фиксаций и комбинированных фиксаций и саккад не показал значимых отличий от результатов первого эксперимента. Рассчитанная погрешность была сравнима с ранее полученными значениями.

Следующий шаг заключался в проведении анализа данных саккад. Значения ошибок первого и второго рода, полученные для описанных данных, представлены в табл. 5.

В ходе анализа была рассмотрена таблица вероятностей для параметра угловой скорости (табл. 6) в целях сравнения полученных значений с первым экспериментом, но значимых отличий не было обнаружено. Наиболее точные результаты предоставляют значения параметров ускоре-

■ **Таблица 5.** Значения ошибок первого и второго рода, %, полученные в ходе анализа саккад взгляда при обработке данных второго эксперимента

Параметр	α-ошибка		β-ошибка	
	Левый глаз	Правый глаз	Левый глаз	Правый глаз
Угловая скорость	50		63,5	
Диаметр зрачка	100	100	0	0
Скорость зрачка	50	50	37,5	25
Скорость взгляда	50	50	12,5	0
Ускорение взгляда	0	0	0	0
Ускорение зрачка	50	50	25	63,5

ния взгляда для левого (табл. 7) и правого (табл. 8) глаза.

На основе представленной таблицы вероятностей можно сформировать следующие кластеры:

$$P_1;$$

$$P_2 = P_3 = P_4 = P_5.$$

Первый кластер является ложным, второй — содержит только одну верную зависимость. Согласно данным кластерам погрешность составляет 75 %, следовательно, данный параметр не обеспечивает необходимой точности идентификации.

Далее были рассмотрены и сравнены значения параметров ускорения взгляда левого и правого глаз. Табл. 7 и 8 предоставляют сравнимые результаты, сформированные на их основе кластеры идентичны:

$$P_1 = P_4;$$

$$P_2 = P_3;$$

$$P_5.$$

■ **Таблица 6.** Таблица вероятностей для параметра угловой скорости саккад взгляда

	P_1	P_2	P_3	P_4	P_5
P_1	1	$3,2 \times 10^{-4}$	$3,7 \times 10^{-3}$	$3,7 \times 10^{-3}$	$1,9 \times 10^{-5}$
P_2		1	0,84	0,81	0,17
P_3			1	0,72	0,43
P_4				1	0,09
P_5					1

■ **Таблица 7.** Таблица вероятностей для параметра ускорения взгляда левого глаза

	P_1	P_2	P_3	P_4	P_5
P_1	1	$9,6 \times 10^{-8}$	$2,5 \times 10^{-7}$	0,11	3×10^{-20}
P_2		1	0,4	3×10^{-7}	$4,5 \times 10^{-6}$
P_3			1	$1,3 \times 10^{-5}$	$1,4 \times 10^{-4}$
P_4				1	$6,5 \times 10^{-18}$
P_5					1

■ **Таблица 8.** Таблица вероятностей для параметра ускорения взгляда правого глаза

	P_1	P_2	P_3	P_4	P_5
P_1	1	$1,4 \times 10^{-9}$	$3,6 \times 10^{-12}$	0,47	$1,9 \times 10^{-22}$
P_2		1	0,06	$6,8 \times 10^{-8}$	$2,5 \times 10^{-6}$
P_3			1	$7,7 \times 10^{-10}$	$1,3 \times 10^{-3}$
P_4				1	6×10^{-20}
P_5					1

■ **Таблица 9.** Таблица результатов проверки нулевой гипотезы теста Колмогорова — Смирнова для параметра ускорения взгляда для данных саккад, полученных в ходе проведения второго эксперимента

	P_1	P_2	P_3	P_4	P_5
P_1	0	1	1	0	1
P_2		0	0	1	1
P_3			0	1	1
P_4				0	1
P_5					0

Согласно представленным кластерам и ключам эксперимента, погрешность составила 0 %, что визуально отображено в таблице результатов теста Колмогорова — Смирнова для проверки нулевой гипотезы (табл. 9). Нулевые значения означают совпадение выборок и, следовательно, положительный результат идентификации.

Анализ данных, полученных в ходе второго эксперимента, показал, что использование параметра ускорения взгляда при саккадах может обеспечить точные результаты, необходимые для осуществления корректной идентификации личности.

Заключение

После завершения экспериментального исследования были достигнуты следующие результаты.

1. Описанный в статье метод показывает, что данные, полученные с помощью систем регистрации траектории взгляда, в том числе и низкочастотных, таких как Tobii X2-30, могут быть использованы для идентификации личности человека.

2. Анализ данных, полученных в ходе проведения двух экспериментов, показал, что ошибка идентификации может варьироваться от 0 до 40 % в зависимости от рассматриваемого параметра движения. В ходе обработки результатов второго эксперимента было выявлено, что анализ саккад обеспечивает более точные результаты по сравнению с анализом параметров, рассчитанных для фиксаций, за счет снятия ограничения на скорости и амплитуды движений глаза.

3. Результаты второго эксперимента показали, что значение ускорения взгляда при рассмотрении саккад может быть использовано для идентификации личности.

Для осуществления идентификации личности с применением анализа ускорения взгляда предполагается, что в режиме реального времени через определенный временной интервал будет производиться расчет значений необходимого параметра на основе полученных с помощью айтрекера данных о траектории взгляда. После чего будет сравниваться полученный результат с записями в базе данных, содержащими рассчитанные значения ускорений взгляда пользователей, и определяться личность человека. Для обеспечения инвариантности системы идентификации относительно зависимости параметров движения глаз от психофизиологического состояния человека в базе данных необходимо хранить одновременно несколько записей, организованных по принципу очереди: новая запись вытесняет самую старую.

Дальнейшие исследования должны быть направлены на разработку новых высокоэффективных алгоритмов идентификации личности человека по характеристикам движения глаз, вычисленных на основе данных низкочастотных систем регистрации траектории взгляда.

Литература

1. Youssif A. A. A., Chowdhury M. U., Ray S., Nafaa H. Y. Fingerprint Recognition System Using Hybrid Matching Techniques // 6th IEEE/ACIS Intern. Conf. on Computer and Information Science. 2007. P. 234–240. doi:10.1109/ICIS.2007.101
2. Nowak P. A Comparative Study on Biometric Hand Identification: Proc. of the 21st Intern. Conf. «Mixed Design of Integrated Circuits & Systems» (MIXDES). 2014. P. 411–414. doi:10.1109/MIXDES.2014.6872231
3. Napieralski J. A., Pastuszka M. M., Sankowski W. 3D Face Geometry Analysis for Biometric Identification: Proc. of the 21st Intern. Conf. «Mixed Design of Integrated Circuits & Systems» (MIXDES). 2014. P. 519–522. doi:10.1109/MIXDES.2014.6872255
4. Ross A., Abaza A. Human Ear Recognition // Computer. 2011. Vol. 44(11). P. 79–81. doi:10.1109/MC.2011.344
5. Павельева Е. А., Крылов А. С., Ушмаев О. С. Развитие информационной технологии идентификации человека по радужной оболочке глаза на основе преобразования Эрмита // Системы высокой доступности. 2009. № 1. С. 36–42.
6. Akram M. U., Tariq A., Khan S. A. Retinal Recognition: Personal Identification Using Blood Vessels // Intern. Conf. for Internet Technology and Secured Transactions (ICITST). 2011. P. 180–184.
7. Сазонов В. А., Садов В. С., Тихоненко С. Г. Идентификация личности по походке на основе скрытых марковских моделей // Электроника инфо. 2012. № 4. С. 105–109.
8. Анисимова Э. Идентификация подписи с помощью радиальных функций // Фундаментальные исследования. 2014. № 9-6. С. 1185–1189.
9. Ходашинский И. А., Савчук М. В., Горбунов И. В., Мещеряков Р. В. Технология усиленной аутенти-

- фикации пользователей информационных процессов // Докл. Томского государственного университета систем управления и радиоэлектроники. 2011. № 2–3. С. 236–248.
10. **Можаров Г. П., Чеботарев Р. С.** Текстонезависимый метод идентификации человека по его голосу // Вестник МГТУ им. Н. Э. Баумана. Сер. Приборостроение. Спец. выпуск «Моделирование и идентификация компьютерных систем и сетей». 2012. С. 168–179.
 11. **Kasprowski P., Ober J.** Enhancing Eye Movement Based Biometric Identification Method by Using Voting Classifiers // Proc. of SPIE. 2005. Vol. 5779. P. 314–324. doi:10.1117/12.603321
 12. **Chan A. D. C., Hamdy M. M., Badre A., Badee V.** Person Identification Using Electrocardiograms // Canadian Conf. on Electrical and Computer Engineering. 2006. P. 1–4. doi:10.1109/CCECE.2006.277291
 13. **Shedeed H. A.** A New Method for Person Identification in a Biometric Security System Based on Brain EEG Signal Processing // 2011 World Congress on Information and Communication Technologies (WICT). 2011. P. 1205–1210. doi: 10.1109/WICT.2011.6141420
 14. **Besbes F., Trichili H., Solaiman B.** Multimodal Biometric System Based on Fingerprint Identification and Iris Recognition // 3rd Intern. Conf. on Information and Communication Technologies: From Theory to Applications. 2008. P. 1–5. doi: 10.1109/ICTTA.2008.4530129
 15. **Kasprowski P., Harężlak K.** The Second Eye Movements Verification and Identification Competition // 2014 IEEE Intern. Joint Conf. on Biometrics (IJCB). 2014. P. 1–6. doi: 10.1109/BTAS.2014.6996285
 16. **Nugrahaningsih N., Porta M., Ricotti S.** Gaze Behavior Analysis in Multiple-Answer Tests: An Eye Tracking Investigation // 2013 Intern. Conf. on Information Technology Based Higher Education and Training (ITHET). 2013. P. 1–6. doi:10.1109/ITHET.2013.6671020
 17. **Pande P., Chandrasekharan S.** Eye-Tracking in STEM Education Research: Limitations, Experiences and Possible Extensions // 2014 IEEE Sixth Intern. Conf. on Technology for Education (T4E). 2014. P. 116–119. doi: 10.1109/T4E.2014.29
 18. **Uskov V., Lyamin A., Lisitsyna L., Sekar B.** Smart e-Learning as a Student-Centered Biotechnical System // Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer, 2014. Vol. 138. P. 167–176. doi: 10.1007/978-3-31913293-8_21
 19. **Lisitsyna L., Lyamin A., Skshidlevsky A.** Estimation of Student Functional State in Learning Management System by Heart Rate Variability Method // Smart Digital Futures 2014. IOS Press, 2014. P. 726–731. doi: 10.3233/978-1-61499-405-3-726
 20. **Лисицына Л. С., Лямин А. В., Быстрицкий А. С., Мартынихин И. А.** Проблема поддержки когнитивных функций в процессе электронного обучения // Научно-технический вестник информационных технологий, механики и оптики. 2014. № 6(94). С. 177–184.
 21. **Лямин А. В., Скушдлевский А. А., Чежин М. С.** Система управления работой компьютерного класса общего доступа // Информационно-управляющие системы. 2011. № 4. С. 66–69.

UDC 004.056.5, 612.846

doi:10.15217/issn1684-8853.2015.5.26

Experimental Studies of Biometric User Identification Based on the Data of Eye Tracker Tobii X2-30

Lyamin A. V.^a, PhD, Tech., Associate Professor, lyamin@mail.ifmo.ru

Cherepovskaya E. N.^a, Senior Assistant, cherepovskaya@cde.ifmo.ru

^aSaint-Petersburg National Research University of Information Technologies, Mechanics and Optics, 49, Kronverkskii St., 197101, Saint-Petersburg, Russian Federation

Purpose: Biometric identification based on eye movement analysis has high protection against forgery. The accuracy of this identification method depends on the analyzed parameters of the eye movement, on the applied stimuli and also on the frequency of the device which records the gaze trajectory. The purpose of our experimental studies is to reveal which eye movement parameters provide the best identification results and which stimuli should be used to reveal personal characteristics of one's eye movement. **Results:** The results of experimental studies of person identification based on eye movement data analysis obtained by eye tracker Tobii X2-30 under two different stimuli are presented. Eye-tracking data obtained by gaze trajectory registration systems, including low-frequency ones, can potentially be used to reveal person's identity. Data analysis for the two experiments showed that the identification error could vary from 0 to 40% over the observed parameter and the gaze event type. The results of the second experiment showed that analyzing saccades' data is more efficient due to the gaze velocity threshold set in order to determine the gaze event type correctly. The second experiment revealed that gaze acceleration parameter of the saccade data could be used to identify a person. **Practical relevance:** The results obtained during the experimental study proved the opportunity of using low-frequency eye-tracking systems for person identification.

Keywords — Information Systems, Biometric Identification Methods, Eye Movement Analysis.

References

1. **Youssif A. A. A., Chowdhury M. U., Ray S., Nafaa H. Y.** Fingerprint Recognition System Using Hybrid Matching Techniques. 6th IEEE/ACIS Intern. Conf. on Computer and Information Science, 2007, pp. 234–240. doi:10.1109/ICIS.2007.101
2. **Nowak P.** A Comparative Study on Biometric Hand Identification. Proc. of the 21st Intern. Conf. Mixed Design of Integrated Circuits & Systems (MIXDES), 2014, pp. 411–414. doi:10.1109/MIXDES.2014.6872231

3. Napieralski J. A., Pastuszka M. M., Sankowski W. 3D Face Geometry Analysis for Biometric Identification. *Proc. of the 21st Intern. Conf. Mixed Design of Integrated Circuits & Systems (MIXDES)*, 2014, pp. 519–522. doi:10.1109/MIXDES.2014.6872255
4. Ross A., Abaza A. Human Ear Recognition. *Computer*, 2011, vol. 44(11), pp. 79–81. doi:10.1109/MC.2011.344
5. Pavelyeva E. A., Krylov A. S., Ushmaev O. S. Development of Human Iris Identification Information Technology Using Hermite Transform. *Sistemy vysokoi dostupnosti*, 2009, no. 1, pp. 36–42 (In Russian).
6. Akram M. U., Tariq A., Khan S. A. Retinal Recognition: Personal Identification Using Blood Vessels. *Intern. Conf. for Internet Technology and Secured Transactions (ICITST)*, 2011, pp. 180–184.
7. Sazonov V. A., Sadov V. S., Tikhonenko S. G. Person identification by Gait Based on Hidden Markov Models. *Elektronika info*, 2012, no. 4, pp. 105–109 (In Russian).
8. Anisimova E. Signature Identification by use of radial Functions. *Fundamental'nye issledovaniia*, 2014, no. 9-6, pp. 1185–1189 (In Russian).
9. Hodashinsky I. A., Savchuk M. V., Gorbunov I. V., Meshcheryakov R. V. Strong Authentication Technology of the Users of Information Processes. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniia i radioelektroniki*, 2011, no. 2–3, pp. 236–248 (In Russian).
10. Mozharov G. P., Chebotarev R. S. Text-Independent Method of Speaker Identification by its Voice. *Vestnik MGTU im. N. E. Bauman. Ser. Priborostroenie. Spetsial'nyi vypusk "Modelirovanie i identifikatsiia komp'iuternykh sistem i setei"*, 2012, pp. 168–179 (In Russian).
11. Kasprowski P., Ober J. Enhancing Eye Movement Based Biometric Identification Method by Using Voting Classifiers. *Proc. of SPIE*, 2005, vol. 5779, pp. 314–324. doi:10.1117/12.603321
12. Chan A. D. C., Hamdy M. M., Badre A., Bader V. Person Identification Using Electrocardiograms. *Canadian Conf. on Electrical and Computer Engineering*, 2006, pp. 1–4. doi:10.1109/CCECE.2006.277291
13. Shedeed H. A. A New Method for Person Identification in a Biometric Security System Based on Brain EEG Signal Processing. *2011 World Congress on Information and Communication Technologies (WICT)*, 2011, pp. 1205–1210. doi:10.1109/WICT.2011.6141420
14. Besbes F., Trichili H., Solaiman B. Multimodal Biometric System Based on Fingerprint Identification and Iris Recognition. *3rd Intern. Conf. on Information and Communication Technologies: From Theory to Applications*, 2008, pp. 1–5. doi:10.1109/ICTTA.2008.4530129
15. Kasprowski P., Harçzłak K. The Second Eye Movements Verification and Identification Competition. *2014 IEEE Intern. Joint Conf. on Biometrics (IJCB)*, 2014, pp. 1–6. doi:10.1109/BTAS.2014.6996285
16. Nugrahaningsih N., Porta M., Ricotti S. Gaze Behavior Analysis in Multiple-Answer Tests: An Eye Tracking Investigation. *2013 Intern. Conf. on Information Technology Based Higher Education and Training (ITHET)*, 2013, pp. 1–6. doi:10.1109/ITHET.2013.6671020
17. Pande P., Chandrasekharan S. Eye-Tracking in STEM Education Research: Limitations, Experiences and Possible Extensions. *2014 IEEE Sixth Intern. Conf. on Technology for Education (T4E)*, 2014, pp. 116–119. doi:10.1109/T4E.2014.29
18. Uskov V., Lyamin A., Lisitsyna L., Sekar B. Smart e-Learning as a Student-Centered Biotechnical System. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Springer, 2014, vol. 138, pp. 167–176. doi:10.1007/978-3-319-13293-8_21
19. Lisitsyna L., Lyamin A., Skshidlevsky A. Estimation of Student Functional State in Learning Management System by Heart Rate Variability Method. *Smart Digital Futures 2014*, IOS Press, 2014, pp. 726–731. doi:10.3233/978-1-61499-405-3-726
20. Lisitsyna L. S., Lyamin A. V., Bystritsky A. S., Martynikhin I. A. Support Problem for Cognitive Functions in the e-Learning. *Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki i optiki*, 2014, no. 6(94), pp. 177–184 (In Russian).
21. Lyamin A. V., Skshidlevsky A. A., Chezhin M. S. Computer-Based Classroom Management System with Public Access. *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2011, no. 4(53), pp. 66–69 (In Russian).

К статьям:

Ben-Shimol Y., Kitroser I. An Efficient Cross-Layer Aware Mapping of VoIP Calls in Wireless OFDMA Systems. Part I. Problem description and channel tracking (*Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2015, no. 2(75), pp. 90–98.).

Ben-Shimol Y., Kitroser I. An Efficient Cross-Layer Aware Mapping of VoIP Calls in Wireless OFDMA Systems. Part II. Mapping Algorithms and their Performance (*Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2015, no. 3(76), pp. 77–84.).

ВМЕСТО

Ben-Shimol Y., PhD, Electrical Engineering, Professor

следует читать

Ben-Shimol Y., PhD, Electrical Engineering, Senior Lecturer

БАЗОВЫЕ МЕТОДЫ И ПРИНЦИПЫ ПОСТРОЕНИЯ АЛГОРИТМОВ ВОССТАНОВЛЕНИЯ РЕГИОНОВ ИЗОБРАЖЕНИЙ

В. А. Ястребов^а, аспирант

А. И. Веселов^а, ассистент

М. Р. Гильмутдинов^а, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Постановка проблемы: в ряде случаев обработка (маскирование) визуальных искажений, возникающих вследствие потерь пакетов данных в сетях с негарантированной доставкой, осуществляется на приемной стороне. Данная задача решается с помощью алгоритмов восстановления изображений. **Цель:** классификация методов и принципов восстановления изображений, выделение достоинств и недостатков существующих подходов, а также актуальных проблем и способов их решения. **Результаты:** рассмотренные подходы и принципы, положенные в основу работы алгоритмов восстановления регионов изображений, разделены на следующие классы: структурные, текстурные и основанные на разреженном представлении данных. Представлены способы борьбы с высокой вычислительной сложностью рассмотренных методов с применением подоптимальных процедур. Отобрана связь между текстурными методами восстановления и методами, основанными на разреженном представлении данных. Приведен анализ качества восстановления различных типов регионов на изображениях. **Практическая значимость:** на базе рассмотренных подходов возможна разработка приложений, критичных к повторной передаче данных по сетям с негарантированной доставкой.

Ключевые слова — восстановление изображений, маскирование визуальных искажений на декодере, разреженное кодирование, структурное восстановление регионов изображений, шаблонное восстановление регионов изображений.

Введение

В настоящее время большое внимание уделяется разработке эффективных методов передачи визуальных данных по сетям связи. Возникновение ошибок или перегрузок при передаче данных по сети приводит к ухудшению качества обрабатываемого изображения. Среди наиболее частых причин потерь пакетов в сети можно выделить:

- перегрузки промежуточных узлов передачи данных;
- получение пакета после требуемого момента его обработки на декодере (актуально для приложений, работающих в режиме реального времени).

Среди существующих методов борьбы с потерями пакетов, возникающими при передаче данных по сети, можно выделить две основные:

- повышение устойчивости передаваемого сигнала к внешним воздействиям (error resilience);
- реконструкцию (маскирование) в целях максимизации степени похожести с оригинальным сигналом (error concealment).

При использовании первого подхода наиболее широко распространенным способом борьбы с ошибками является применение помехоустойчивого кодирования. Если возникшие ошибки не удастся полностью исправить, то используется ретрансляция данных, т. е. повторное отправленное источником данных, не дошедших до приемника. При этом необходимым условием является

существование обратной связи между передатчиком (кодером) и получателем (декодером) данных. Однако такой подход не всегда практически осуществим, особенно в системах, которые критичны к повторным передачам, ввиду увеличения задержки получения данных. Кроме того, в некоторых случаях обратной связи не существует, или по ряду причин воспользоваться ею не представляется возможным.

В связи с этим актуальной задачей является обработка вызванных такими потерями искажений только на приемной стороне, т. е. так называемая процедура маскирования искажений.

Исследованию методов и принципов восстановления изображений и посвящена данная работа. В основу рассмотренной классификации положены определяющие принципы построения алгоритмов восстановления регионов изображений, изложенные в базовых работах [1–4]. Дальнейшие публикации [5–7] подтверждают актуальность изложенных принципов.

Постановка задачи восстановления регионов изображений

Рассмотрим задачу восстановления регионов изображений на стороне приемника на примере типовой системы передачи данных. На основе блочной обработки изображения X осуществляется формирование сжатого битового потока, который впоследствии разбивается на пакеты в целях дальнейшей передачи по сети. Будем счи-

тать, что в процессе передачи произошло некоторое событие, в результате которого часть пакетов не была получена приемником, из-за чего он не смог восстановить на изображении некоторые области. Назовем подобные области искаженными регионами. Обозначим множество пикселей в искаженных регионах через Ω , а множество корректно полученных пикселей через $\bar{\Omega}$. Процесс восстановления значений интенсивностей пикселей из Ω при известных значениях пикселей из $\bar{\Omega}$ будем называть маскированием искажений. Обозначим пиксель изображения X , находящийся на пересечении i -й строки и j -го столбца, как $X(i, j)$, а квадратную область размером $M \times M$ (блок) на изображении X , центральный пиксель которого имеет координату (i, j) , как $B(i, j)$.

Введем ряд допущений применительно к рассматриваемой системе обработки и передачи данных (рис. 1):

- 1) приемник всегда может определить факт потери пакета;
- 2) блоки на стороне передатчика кодируются независимо;
- 3) в каждом пакете передается информация только об одном блоке;
- 4) искаженные блоки не могут находиться на краях изображения.

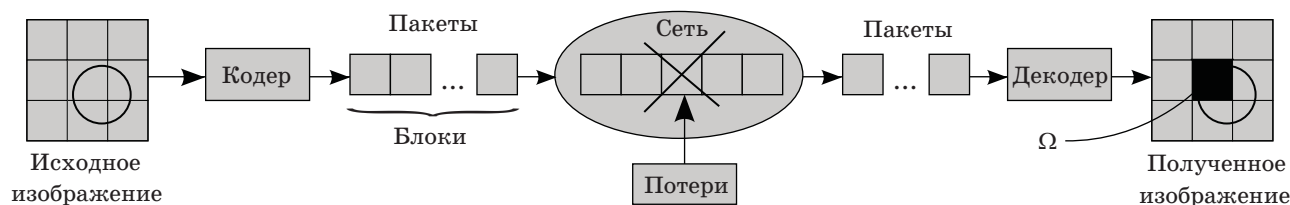
Будем считать, что потеря пакета приводит к появлению черного квадрата по координатам, соответствующим координатам потерянного блока.

Следует отметить, что приведенные допущения не отражают реальную природу визуальных искажений в восстановленном на приемнике изображении, так как, как правило, блоки кодируются независимо, и потеря пакета на практике может привести к искажению нескольких блоков в любой позиции изображения. Однако приведенные выше допущения вполне достаточны для рассмотрения базовых принципов и методов восстановления искаженных регионов на изображениях, что и является основной целью данной работы.

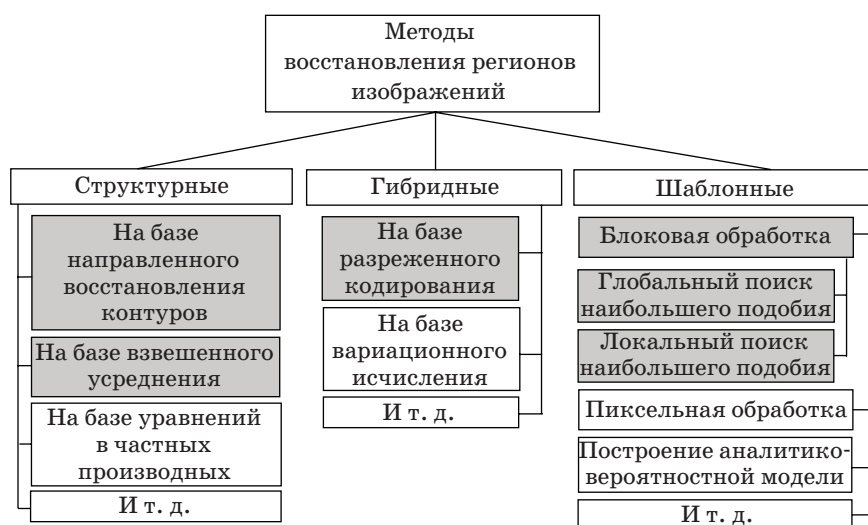
Рассмотрим базовые принципы восстановления изображений [1].

1. Данные, принимаемые для маскирования, должны определяться всем изображением X в целом. Восстановление данных внутри искаженного региона Ω должно производиться данными, являющимися естественным продолжением информации, содержащейся в доступном регионе $\bar{\Omega}$.

2. Структура синтезированных для восстановления данных должна отражать структуру неискаженных данных. Как правило, это достигается путем интерполяции контуров объектов, обнаруженных на границе $\partial\Omega$. Обозначим через $\partial\Omega$



■ Рис. 1. Рассматриваемая система обработки и передачи данных



■ Рис. 2. Классификация методов восстановления изображений

подмножество пикселей из $\bar{\Omega}$, смежных с пикселями из Ω , а через $\partial\bar{\Omega}$ подмножество пикселей из Ω , смежных с пикселями из $\bar{\Omega}$. При этом конкретный состав этих подмножеств зависит от используемого алгоритма восстановления.

3. Цветовая составляющая восстановленного региона должна определяться анализом интенсивностей пикселей, окружающих искаженный регион Ω .

4. Восстановленные данные должны учитывать локальные особенности обрабатываемого изображения X , например текстуру объектов.

В зависимости от того, как указанные принципы учитываются при восстановлении искаженных регионов, каждый класс применяет свой метод восстановления: структурный, шаблонный или гибридный.

Методы восстановления изображений, рассматриваемые далее, на рис. 2 выделены серым цветом.

Структурные методы восстановления

Использование структурных методов для восстановления регионов изображений подразумевает применение процедуры, которая опирается на некоторую априорную информацию о структуре данных в искаженном блоке. Как правило, для определения априорной информации выполняется анализ блоков, смежных с искаженным. В результате искаженные блоки относят к одному из двух типов:

- блоки гладкого типа;
- блоки, содержащие контуры объектов.

Данная классификация блоков по типам обусловлена тем, что визуальное восприятие человека крайне чувствительно к контурам объектов [8]. Поэтому одной из важных задач любого алгоритма восстановления визуальных данных является как точное восстановление нарушенных контуров, так и минимизация вероятности формирования ложных контуров, отсутствующих на исходном изображении. Кроме того, восстановление границ объектов в расплывчатом или неявном виде также крайне негативно сказывается на восприятии визуальных данных в целом.

Структурные методы восстановления можно разделить на два подкласса:

- 1) на основе взвешенного усреднения значений пикселей (weighted averaging);
- 2) на основе направленного восстановления контуров (directional interpolation).

Первые, как правило, применяются для восстановления блоков гладкого типа. Например, в работе [9] предлагается производить восстановление искаженных пикселей путем взвешенного усреднения значений пикселей на границе $\partial\bar{\Omega}$. Значения весовых коэффициентов обратно про-

порциональны расстоянию между интерполируемым пикселем $X(k, r)$ из искаженного блока $B(i, j)$ и пикселями, значения интенсивности которых известны:

$$X(k, r) = \frac{(\omega_R X(k, j-1) + \omega_L X(k, j+M) + \omega_T X(i+M, r) + \omega_B X(i-1, r))}{\omega_L + \omega_R + \omega_T + \omega_B},$$

где ω_R , ω_L , ω_T и ω_B — расстояния между интерполируемым пикселем и ближайшими пикселями на границе известных соседних блоков; $X(i-1, r)$, $X(k, j-1)$, $X(k, j+M)$ и $X(i+M, r)$ — значения ближайших пикселей, расположенных на границе потерянного блока по четырем направлениям; $M \times M$ — размер потерянного блока.

В большинстве методов, где применяется восстановление блоков, содержащих контуры, можно выделить три основных этапа:

- 1) оценка направления контуров в искаженном блоке за счет анализа пикселей из неискаженных блоков;
- 2) восстановление контуров на основе информации, содержащейся в соседних блоках данных региона Ω ;
- 3) восстановление оставшихся пикселей изображения, а также постобработка всего восстановленного региона целиком.

В работах [10, 11] предлагается восстанавливать контуры, содержащие разрывы, используя билинейную интерполяцию, опирающуюся на данные об уже обнаруженных контурах в соседних блоках. Алгоритм предполагает разбиение потерянного блока $B(i, j)$ на несколько регионов в зависимости от произведенной оценки направления контура, при этом предлагается проводить такое разбиение, которое учитывает как направление доминирующего контура, так и направления возмущенных менее ярко выраженных контуров.

Таким образом, алгоритмы на базе взвешенного усреднения способны хорошо восстанавливать гладкие области на изображении, а методы на базе направленного восстановления контуров — области, содержащие контуры.

Шаблонные методы восстановления

Основная идея работы алгоритмов этого класса основана на предположении о наличии повторяющихся фрагментов данных на изображении, которые в дальнейшем будем называть шаблонами. Восстановление целевого региона Ω производится по частям путем копирования значений пикселей из наиболее похожего шаблона. Как правило, в методах этого класса можно выделить два основных этапа. На первом осуществляется расстановка приоритетов для пикселей $X(i, j)$ из $\partial\Omega$. Затем вокруг пикселя с наибольшим при-

оритетом строится блок $B(i, j)$. На втором проводится поиск наиболее похожего блока $B(k, l)$, $k \neq i$, $l \neq j$, из доступного набора данных Ω на основе выбранного критерия схожести, например, суммы абсолютных разностей или суммы квадратов разностей [12] (вычисляемой только для пикселей из Ω). Восстановление пикселей из $B(i, j) \cap \Omega$ выполняется путем копирования значений пикселей из найденного блока $B(k, l)$ на соответствующих позициях. Описанная процедура итеративно повторяется до тех пор, пока целевая область Ω не будет восстановлена полностью.

Глобальный алгоритм поиска, основанный на принципе наибольшей окрестной схожести (best neighborhood matching), был впервые предложен в работе [2]. Поскольку рассмотренные далее в данном разделе алгоритмы являются его модификациями, то приведем шаги этого алгоритма подробнее.

1. Инициализация искаженной (целевой) области Ω .

2. Расстановка приоритетов восстановления пикселей из целевой области Ω . В начале каждой новой итерации алгоритма производится расчет порядка заполнения целевых пикселей с помощью применения заданной функции расстановки приоритетов $P(X(i, j))$.

3. Поиск шаблона и интерполяция. Поврежденный блок данных извлекается вместе с окружающими его известными пикселями из Ω . Предлагается использовать глобальный поиск для анализа всех блоков-кандидатов из Ω .

4. Обновление списка целевых пикселей, а также прочей информации, требуемой для выставления приоритетов целевым пикселям.

Недостатком алгоритма [2] является использование глобального поиска, имеющего высокую вычислительную сложность. В связи с этим в работе [13] рассмотрены ключевые параметры, используемые в подоптимальном поиске: процедура поиска и порядок восстановления пикселей, которые влияют как на качество восстановления, так и на скорость работы алгоритма.

В работах [14–16] рассмотрено применение подоптимальных алгоритмов для ускорения работы. В работах [14, 15] предложена модификация методики поиска на основе применения идеи градиентного спуска. Обозначим квадратный регион на изображении X , центральным элементом которого является пиксель $X(i, j)$, как окно поиска $N(X(i, j))$. В работе [15] выбор области поиска уменьшается путем установки шага сдвига окна поиска $N(X(i, j))$ в значение, превышающее один пиксель. В работе [14] предложено осуществлять выбор на основе анализа соседних регионов, используя спиралевидную траекторию поиска. В работе [16] предложено производить поиск в компонентах, полученных в результате вейвлет-преобразования.

Поскольку вопрос расстановки приоритетов для пикселей является специфическим для задачи восстановления регионов изображений, то рассмотрим его подробнее.

Искаженный регион изображения Ω может содержать как текстурную, так и структурную информацию. В статье [3] было показано, что для достижения высокого качества восстановления алгоритм должен уметь находить и разделять эти составляющие в целях приоритетного восстановления именно структурной части потерянных данных. В результате проведенных исследований авторы предложили подход, восстанавливающий в первую очередь данные, характеризующие именно структуру потерянного региона Ω . Для каждого пикселя $X(i, j)$ на границе $\partial\Omega$ выставляется приоритет восстановления путем анализа блока, центральным элементом которого является пиксель $X(i, j)$. Функция расстановки приоритетов определена следующим образом:

$$P(X(i, j)) = C(X(i, j))D(X(i, j)).$$

Первый сомножитель — доверительный (confidence term), задает степень уверенности в восстановленных данных путем анализа соотношения количества известных пикселей к количеству неизвестных в рассматриваемом блоке. Второй сомножитель — информационный (data term), отражает количество структурных данных, содержащихся в обрабатываемом блоке. Значение информационного сомножителя зависит от изофоты (силы контура), пересекающей границу $\partial\Omega$ неизвестного региона Ω .

Благодаря описанной функции расстановки приоритетов в первую очередь осуществляется интерполяция пикселей на границах объектов, а только затем в однородных блоках данных.

Согласно исследованиям [17], описанный алгоритм обладает одним существенным недостатком. В процессе заполнения целевых пикселей региона Ω степень уверенности в интерполируемых данных неуклонно падает. Это выражается в том, что значение доверительного сомножителя $C(X(i, j))$ в функции расстановки приоритетов $P(X(i, j))$ с ростом количества итераций стремится к нулю, способность же роста значения информационного сомножителя $D(X(i, j))$ сильно ограничена. Эти факторы могут привести к неправильному порядку восстановления целевых пикселей региона Ω , поскольку в этом случае результат, выдаваемый функцией $P(X(i, j))$, перестает зависеть от входных данных.

Для устранения этого недостатка в работе [18] была предложена модифицированная функция расстановки приоритетов, в которой умножение было заменено сложением, что позволило сделать ее более устойчивой к шумовым воздействиям. Авторами также модифицирована

функция расчета уверенности путем оценивания степени размытости кривой γ : $P(B(i, j)) = (1 - \gamma) \cdot C(B(i, j)) + \gamma$. В итоге функция выставления приоритетов определена следующим образом:

$$P(B(i, j)) = \alpha C(B(i, j)) + (1 - \alpha) D(B(i, j)),$$

где $0 \leq \alpha \leq 1$ — весовой коэффициент.

Шаблонные методы восстановления способны показывать хорошие результаты при обработке текстурных областей на изображении, поскольку те содержат множество повторяющихся фрагментов. Однако они не предоставляют наилучшего способа обработки регионов, содержащих большое количество контуров объектов, которые лучше восстанавливаются структурными методами. В связи с этим ряд исследований посвящен разработке алгоритмов, позволяющих производить эффективное восстановление как текстурных, так и структурных регионов изображения. Такие алгоритмы формируют класс гибридных методов восстановления. Одно из наиболее перспективных направлений в этом классе основывается на использовании разреженного кодирования для восстановления регионов изображений. В связи с этим в данной статье центральное внимание будет уделено гибридным методам восстановления, основанным на разреженном представлении данных.

Методы восстановления, основанные на разреженном кодировании данных

Главная идея разреженного кодирования заключается в представлении исходного сигнала x размерности N другим сигналом α размерности K , где $K > N$. Представление сигнала α при этом должно содержать малое количество ненулевых коэффициентов.

Разреженное представление сигнала можно получить с помощью разложения исходного сигнала по расширенному базису, который принято называть словарем. Словарь содержит все базисные векторы, а также некоторое число их линейных комбинаций, т. е. число векторов, содержащихся в нем, превышает размерность исходного сигнала. Именно благодаря такому расширенному базису можно получить разреженное представление сигнала.

Обозначим одномерное представление (как правило, в виде вектора-столбца) конкретного блока с центральным пикселем $X(i, j)$ изображения X как $\Psi(i, j)$. Далее будем называть $\Psi(i, j)$ патчем. Таким образом, соотношение между патчем $\Psi(i, j)$ и блоком $B(i, j)$ можно выразить следующим равенством:

$$\Psi(i, j) = \phi(B(i, j)),$$

где ϕ — функция обхода блока в сканирующем порядке.

Существует два основных подхода, которые применяют методы разреженного кодирования данных для построения алгоритмов восстановления изображений:

— синтез потерянного региона Ω на основе анализа линейных зависимостей, обнаруженных на изображении [19];

— использование словарей для восстановления патчей изображения (словарный метод) [4].

Авторы работы [19] предлагают использовать принцип, который основан на предположении о том, что потерянный патч может быть представлен линейной комбинацией нескольких других патчей, сформированных из $\bar{\Omega}$. Для того чтобы определить, содержит ли восстанавливаемый патч текстурную или структурную информацию, было предложено анализировать количество соседних патчей $\Psi(i, j)$, необходимых для его аппроксимации. При этом было установлено, что структурные патчи имеют более разреженную степень представления (требуют меньшее количество патчей-кандидатов для точного представления) по сравнению с текстурными патчами.

Предлагается [19] в первую очередь восстанавливать структурные патчи. По аналогии с шаблонными методами восстановления алгоритм состоит из двух основных шагов (рис. 3):

1) расстановка приоритетов восстановления для патчей;

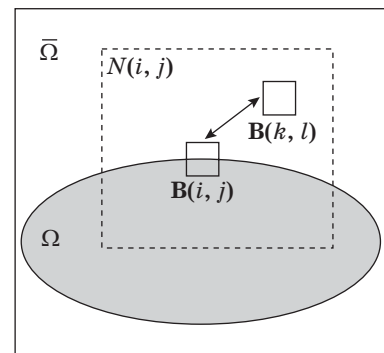
2) восстановление искаженных патчей.

Рассмотрим процедуру расчета приоритета для произвольного патча $\Psi(i, j)$.

Для этого определим множество патчей-кандидатов $N(i, j)$:

$$N(i, j) = \{ \Psi(k, l) : |k - i| < r, |l - j| < r, \Psi(k, l) \subset \bar{\Omega} \},$$

где r — некоторая заранее определенная константа (радиус). Таким образом, множество $N(i, j)$ содержит патчи из неискаженной области $\bar{\Omega}$, состоящие не более чем на r пикселей по горизонтали и вертикали от координаты (i, j) .



■ Рис. 3. Процедура поиска блока-кандидата

Функция расстановки приоритетов выглядит следующим образом:

$$P(\Psi(i, j)) = T_{[\zeta, 1]} \left(\left\| \eta_{(i, j)} \right\|_2 \sqrt{\frac{|N(i, j) \cdot M^2|}{(2r + 1)^2}} C(\Psi(i, j)) \right),$$

где $T_{[\zeta, 1]}$ — функция приведения числовых данных в требуемый диапазон $[\zeta, 1]$; $\|\cdot\|_2$ — L2-норма вектора; $\eta_{(i, j)}$ содержит весовые коэффициенты, характеризующие степень схожести между исходным патчем $\Psi(i, j)$ и патчами-кандидатами $\Psi(k, l) \in N(i, j)$; $|\cdot|$ — мощность множества; M^2 — размер патча; $C(\Psi(i, j))$ — функция выставления уверенности (по аналогии с шаблонными методами восстановления).

Восстановление искаженного патча $\Psi(i, j)$ достигается путем построения линейной комбинации из нескольких наиболее похожих патчей, найденных на изображении. Вопросы, связанные с влиянием количества патчей для аппроксимации, а также выбор способа расчета вектора $\eta_{(i, j)}$ детально рассмотрены в работе [19].

Таким образом, высокое значение функции $P(\Psi(i, j))$ указывает на то, что рассматриваемый патч является структурным, поскольку его представление сильно разрежено. Текстурные же патчи таким свойством не обладают, что приводит к низким значениям функции $P(\Psi(i, j))$.

Перейдем к методам восстановления на основе использования функции для аппроксимации изображения. В качестве примера рассмотрим словарный метод [4], в котором разреженное кодирование применяется для выделения слоев на изображении, содержащих текстурные и структурные элементы. Основная идея данного подхода заключается в использовании двух заранее подготовленных словарей, один из которых позволяет получить разреженное представление текстурной информации, а другой — структурной. Такой подход принято называть морфологическим анализом компонент (morphological component analysis). В таком случае изображение рассматривается как поэлементная сумма двух составляющих (морфологических слоев):

$$\mathbf{X} = \mathbf{X}_t + \mathbf{X}_s,$$

где $\mathbf{X}_t$ — текстурная, а $\mathbf{X}_s$ — структурная составляющая. Для упрощения дальнейшего изложения будем полагать, что $\mathbf{X}$, $\mathbf{X}_t$ и $\mathbf{X}_s$ представлены в виде векторов в пространстве $\mathbb{R}^n$, где n — величина, равная произведению ширины исходного изображения на его длину.

Поскольку в ходе работы алгоритма морфологические компоненты представляются в разреженном виде, то окончательное изображение $\mathbf{X}$ строится по следующему принципу:

$$\mathbf{X} = \mathbf{T}_t \alpha_t^{\text{opt}} + \mathbf{T}_s \alpha_s^{\text{opt}},$$

где $\mathbf{T}_t, \mathbf{T}_s$ — словарь, позволяющий получить разреженное представление текстурного слоя $\mathbf{X}_t$ и структурной компоненты изображения $\mathbf{X}_s$ соответственно.

Поиск оптимальных коэффициентов $\{\alpha_t^{\text{opt}}, \alpha_s^{\text{opt}}\}$ предлагается осуществлять путем решения следующей оптимизационной задачи:

$$\{\alpha_t^{\text{opt}}, \alpha_s^{\text{opt}}\} = \arg \min_{\{\alpha_t, \alpha_s\}} \left\{ \left\| \alpha_t \right\|_0 + \left\| \alpha_s \right\|_0 + \lambda \left\| \mathbf{V}^T (\mathbf{X} - \mathbf{T}_t \alpha_t - \mathbf{T}_s \alpha_s) \right\|_2^2 + \gamma TV\{\mathbf{T}_s \alpha_s\} \right\},$$

где α_t^{opt} и α_s^{opt} принадлежат пространству $\mathbb{R}^L$ ($L > M^2$); $\mathbf{V}$ — вектор, содержащий единицы на позициях с доступными пикселями, а нули — на позициях с искаженными, принадлежит пространству $\mathbb{R}^n$; λ и γ — весовые коэффициенты; TV — дополнительные ограничения, способствующие кусочно-гладкому восстановлению структурного слоя [20]. Решение данной оптимизационной задачи можно находить, используя метод преследования базисов (basis pursuit) [21].

Хотелось бы отметить, что методы, основанные на разреженном представлении данных, во многом схожи с шаблонными методами по принципу обработки отдельных блоков, взятых с изображения. Однако алгоритмы, применяющие шаблонные методы восстановления, в процессе своей работы не создают новых шаблонов, а используют только уже существующие. Алгоритмы восстановления регионов изображений, основанные на разреженном представлении данных, за счет комбинирования базисных векторов словаря способны создавать новые шаблоны. В связи с этим, как будет показано ниже, в ряде случаев они способны показывать более хорошие результаты при восстановлении искаженных регионов.

Анализ качества восстановления различных типов регионов на изображениях

Результат анализа эффективности восстановления различных типов регионов на изображениях с помощью рассмотренных в данной работе методов представлен в таблице.

Примеры типов рассматриваемых регионов изображений приведены на рис. 4.

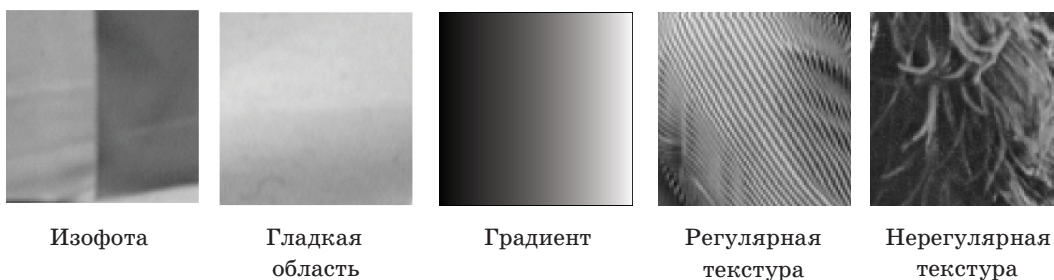
В таблице приняты следующие условные обозначения:

+ : в абсолютном большинстве случаев восстановленное изображение не содержит серьезных визуальных искажений;

+? : в большинстве случаев восстановленное изображение не содержит серьезных визуальных искажений;

■ Анализ качества восстановления различных типов регионов на изображениях

Метод восстановления	Тип региона				
	Изофота	Гладкая область	Градиент	Регулярная текстура	Нерегулярная текстура
Структурные методы: взвешенное усреднение направленное восстановление контуров	–?	+	+	–	–
	+	+	+	–?	–
Шаблонные методы: глобальный поиск локальный поиск	–?	+	+?	+	+
	–?	+	+?	+	+?
Разреженное кодирование	+?	+	+	+	–?



■ Рис. 4. Примеры типов регионов на изображении

–?: в некоторых случаях восстановленное изображение не содержит серьезных визуальных искажений;

–: в абсолютном большинстве случаев восстановленное изображение может содержать заметные визуальные искажения.

Алгоритмы восстановления, основанные на взвешенном усреднении, способны качественно восстанавливать области изображений, не содержащие каких-либо особенностей. К таким регионам можно отнести гладкие области, а также области, содержащие градиенты. В отличие от них методы направленного восстановления контуров способны качественно восстанавливать и регионы, содержащие внутри себя незначительные вариации. Таким образом, помимо вышеперечисленных регионов они способны корректно обрабатывать и области, содержащие изофоты, и регулярные текстуры, если те в свою очередь состоят из изофот.

Алгоритмы, основанные на применении глобального или локального поиска, осуществляют восстановление путем использования готовых шаблонов данных на изображении. Принцип их работы не предусматривает модификацию или изменение таких шаблонов. Таким образом, они способны корректно восстанавливать потерянные регионы, содержащие информацию, которая обладает значительной корреляцией с той, что содержится в доступной области Ω .

Алгоритмы на основе разреженного представления данных для восстановления используют свойство словарей «запоминать» повторяющиеся

особенности данных на изображении. При этом такие алгоритмы способны генерировать блоки данных, содержащие информацию, отсутствующую в явном виде в доступной области Ω . Таким образом, самыми сложными регионами для восстановления будут те, что содержат нерегулярную текстуру.

Закключение

В работе выделены и проанализированы классы алгоритмов восстановления изображений, которые основаны на структурных и шаблонных методах, а также на разреженном кодировании данных.

Проведен анализ эффективности восстановления различных типов регионов на изображениях с помощью рассмотренных алгоритмов, в результате которого определены достоинства и недостатки рассмотренных подходов.

Рассмотренные методы восстановления изображений исключительно на стороне получателя данных могут быть применены в задаче маскирования искажений, возникающих в результате потерь пакетов в сети передачи данных. Анализ и разработка методов маскирования искажений визуальных данных является актуальной задачей для различных прикладных областей и требует дальнейших исследований.

Работа выполнена при финансовой поддержке Минобрнауки в рамках базовой части государственного задания вузу на 2015 год (проект № 2452).

Литература

1. Bertalmio M., Sapiro G., Caselles V., Ballester C. Image Inpainting // Proc. of the 27th Annual Conf. on Computer Graphics and Interactive Techniques. 2000. P. 417–424.
2. Wang Z., Yu Y., Zhang D. Best Neighborhood Matching: an Information Loss Restoration Technique for Block-Based Image Coding Systems // IEEE Transactions on Image Processing. 1998. N 7. P. 1056–1061.
3. Criminisi A., Perez P., Toyama K. Region Filling and Object Removal by Exemplar-Based Inpainting // IEEE Transactions on Image Processing. 2004. N 13. P. 1200–1212.
4. Elad M., Starck J.-L., Querre P., Donoho D. Simultaneous Cartoon and Texture Image Inpainting Using Morphological Component Analysis (MCA) // Applied and Computational Harmonic Analysis. 2005. N 19. P. 340–358.
5. Qaratlu M. M., Ghanbari M. Intra-frame Loss Concealment Based on Directional Extrapolation // Signal Processing: Image Communication. 2011. N 26. P. 304–309.
6. Xiang Y., Feng L., Xie S., Zhou Z. An Efficient Spatio-Temporal Boundary Matching Algorithm for Video Error Concealment // Multimedia Tools and Applications. 2011. N 52. P. 91–103.
7. Nguen D., Dao M., Tran T. Video Error Concealment Using Sparse Recovery and Local Dictionaries // Processing of ICASSP. 2011. P. 1125–1128.
8. Chan T. F., Shen J. Mathematical Models for Local Nontexture Inpaintings // SIAM J. Appl. Math. 2002. P. 1019–1043.
9. Salama P., Shroff N. B., Delp E. J. Error Concealment in Encoded Video Streams // Signal Recovery Techniques for Image and Video Compression and Transmission 1998. P. 119–233.
10. Nemethova O., Al-Moghrabi A., Rupp M. Flexible Error Concealment for H.264 Based on Directional Interpolation // Proc. of Intern. Conf. on Wireless Networks, Communications and Mobile Computing. 2005. P. 1255–1260.
11. Kung W.-Y., Kim C.-S., Kuo C. C. Spatial and Temporal Error Concealment Techniques for Video Transmission over Noisy Channels // IEEE Transactions on Circuits and Systems for Video Technology. 2006. N 16. P. 789–803.
12. Ulysses J. N., Conci A. Measuring Similarity in Medical Registration // 17th Intern. Conf. on Systems, Signals and Image Processing. 2010. http://www2.ic.uff.br/iwssip2010/Proceedings/nav/papers/paper_189.pdf (дата обращения: 13.07.2015).
13. Zhang D., Wang Z. Image Information Restoration Based on Long-Range Correlation // IEEE Transactions on Circuits and Systems for Video Technology. 2002. N 12. P. 331–341.
14. He L., Zhang Y. A Rotate-Based Best Neighborhood Matching Algorithm for High Definition Image Error Concealment // Proc. of the 2010 10th IEEE Intern. Conf. on Computer and Information Technology. 2010. P. 1393–1396.
15. Li W., Zhang D., Liu Z., Qiao X. Fast Block-Based Image Restoration Employing the Improved Best Neighborhood Matching Approach // IEEE Transactions on Systems, Man, and Cybernetics. 2005. N 35. P. 546–555.
16. Hsieh C.-T., Chen Y.-L., Hsu C.-H. Fast Image Restoration Using the Multi-Layer Best Neighborhood Matching Approach // Proc. of the 6th WSEAS Intern. Conf. on Applied Computer Science. 2007. P. 441–446.
17. Nalawade V. V., Ruikar S. D. A Comparative Analysis of Exemplar Based and Wavelet Based Inpainting Technique // Intern. Journal of Electronics and Computer Science Engineering. 2012. N 1. P. 1034–1043.
18. Cheng W., Hsieh C., Lin S., Wang C., Wu J. Robust Algorithm for Exemplar Based Image Inpainting // Proc. of Intern. Conf. on Computer Graphics, Imaging and Visualization. 2005. P. 64–69.
19. Xu Z., Sun J. Image Inpainting by Patch Propagation Using Patch Sparsity // IEEE Transactions on Image Processing. 2010. N 19. P. 1153–1165.
20. Rudin L., Osher S., Fatemi E. Nonlinear Total Variation Noise Removal Algorithm // Physica D. 1992. N 60. P. 259–268.
21. Chen S. S., Donoho D. L., Saunders M. A. Atomic Decomposition by Basis Pursuit // SIAM Journal on Scientific Computing. 1998. N 20. P. 33–61.

UDC 004.932.2

doi:10.15217/issn1684-8853.2015.5.34

Basic Methods and Principles of Developing Image Inpainting Algorithms

Yastrebov V. A.^a, Post-Graduate Student, victor.yastrebov1@yandex.ruVeselov A. I.^a, Assistant Professor, anton.veselov@gmail.ruGilmudtinov M. R.^a, PhD, Tech., Associate Professor, mgilm@gmail.ru^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., Saint-Petersburg, 190000, Russian Federation

Purpose: Visual distortions caused by lost packets in networks with non-guaranteed delivery are sometimes processed (masked) at the decoder side. To solve this problem, image inpainting algorithms are used. The purpose of this work is classifying the methods

and principles of these algorithms, finding the advantages and shortcomings of the existing approaches, formulating the current problems and possible ways to resolve them. **Results:** The discussed approaches and principles which form the basis for image inpainting algorithms are divided into the following classes: structural techniques, textural techniques and sparse-coding techniques. Some ways are presented to overcome the high computational complexity of the discussed methods using suboptimal procedures. The link is shown between textural inpainting techniques and sparse-coding techniques. Restoration quality analysis is given for various image region types. **Practical relevance:** The discussed approaches can be applied for the development of applications sensitive to retransmission in networks with non-guaranteed delivery.

Keywords — Image Reconstruction, Visual Distortion Masking, Sparse Coding Inpainting, Structure Inpainting, Exemplar-based Inpainting.

References

1. Bertalmio M., Sapiro G., Caselles V., Ballester C. Image Inpainting. *Proc. of the 27th Annual Conf. on Computer Graphics and Interactive Techniques*, 2000, pp. 417–424.
2. Wang Z., Yu Y., Zhang D. Best Neighborhood Matching: an Information Loss Restoration Technique for Block-Based Image Coding Systems. *IEEE Transactions on Image Processing*, 1998, vol. 7, pp. 1056–1061.
3. Criminisi A., Perez P., Toyama K. Region Filling and Object Removal by Exemplar-Based Inpainting. *IEEE Transactions on Image Processing*, 2004, vol. 13, pp. 1200–1212.
4. Elad M., Starck J.-L., Querre P., Donoho D. Simultaneous Cartoon and Texture Image Inpainting using Morphological Component Analysis (MCA) *Applied and Computational Harmonic Analysis*, 2005, no. 19(3), pp. 340–358.
5. Qaratlu M. M., Ghanbari M. Intra-Frame Loss Concealment Based on Directional Extrapolation. *Signal Processing: Image Communication*, 2011, no. 26, pp. 304–309.
6. Xiang Y., Feng L., Xie S., Zhou Z. An Efficient Spatio-Temporal Boundary Matching Algorithm for Video Error Concealment. *Multimedia Tools and Applications*, 2011, no. 52, pp. 91–103.
7. Nguen D., Dao M., Tran T. Video Error Concealment Using Sparse Recovery and Local Dictionaries. *Processing of ICASSP*, 2011, pp. 1125–1128.
8. Chan T. F., Shen J. Mathematical Models for Local Nontexture Inpaintings. *SIAM J. Appl. Math.*, 2002, no. 62, pp. 1019–1043.
9. Salama P., Shroff N. B., Delp E. J. Error Concealment in Encoded Video Streams. *Signal Recovery Techniques for Image and Video Compression and Transmission*, 1998, pp. 119–233.
10. Nemethova O., Al-Moghrabi A., Rupp M. Flexible Error Concealment for H.264 Based on Directional Interpolation. *Proc. of Intern. Conf. on Wireless Networks, Communications and Mobile Computing*, 2005, pp. 1255–1260.
11. Kung W.-Y., Kim C.-S., Kuo C. C. Spatial and Temporal Error Concealment Techniques for Video Transmission Over Noisy Channels. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, no. 16(7), pp. 789–803.
12. Ulysses J. N., Conci A. Measuring Similarity in Medical Registration. *17th Intern. Conf. on Systems, Signals and Image Processing*, 2010. Available at http://www2.ic.uff.br/iwssip2010/Proceedings/nav/papers/paper_189.pdf (accessed 13 July 2015).
13. Zhang D., Wang Z. Image Information Restoration Based on Long-Range Correlation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2002, vol. 12, pp. 331–341.
14. He L., Zhang Y. A Rotate-Based Best Neighborhood Matching Algorithm. *Intern. Conf. on Computer and Information Technology*, 2010, pp. 1393–1396.
15. Li W., Zhang D., Liu Z., Qiao X. Fast Block-Based Image Restoration Employing the Improved Best Neighborhood Matching Approach. *IEEE Transactions on Systems, Man, and Cybernetics*, 2005, no. 35(4), pp. 546–555.
16. Hsieh C.-T., Chen Y.-L., Hsu C.-H. Fast Image Restoration Using the Multi-Layer Best Neighborhood Matching Approach. *Proc. of the 6th WSEAS Intern. Conf. on Applied Computer Science*, 2007, pp. 441–446.
17. Nalawade V. V., Ruikar S. D. A Comparative Analysis of Exemplar Based and Wavelet Based Inpainting Technique. *Intern. Journal of Electronics and Computer Science Engineering*, 2012, no. 1(3), pp. 1034–1043.
18. Cheng W., Hsieh C., Lin S., Wang C., Wu J. Robust Algorithm for Exemplar Based Image Inpainting. *Proc. of Intern. Conf. on Computer Graphics, Imaging and Visualization*, 2005, pp. 64–69.
19. Xu Z., Sun J. Image Inpainting by Patch Propagation Using Patch Sparsity. *IEEE Transactions on Image Processing*, 2010, no. 19(5), pp. 1153–1165.
20. Rudin L., Osher S., Fatemi E. Nonlinear Total Variation Noise Removal Algorithm. *Physica D*, 1992, no. 60(1), pp. 259–268.
21. Chen S. S., Donoho D. L., Saunders M. A. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 1998, no. 20, pp. 33–61.

ЛАЗЕРНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ: ПРИНЦИПЫ ПОСТРОЕНИЯ СТАТИСТИЧЕСКОЙ ТЕОРИИ

В. И. Хименко^а, доктор техн. наук, профессор

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Цель: обобщение и развитие результатов по формированию статистической радиооптики — единой статистической теории лазерных информационных систем и радиооптических методов получения, преобразования, передачи и обработки информации. **Результаты:** выделены основные трудности и показан противоречивый характер развития радиооптики. Предложена обобщенная модель радиооптической обработки информации и обоснован общий принцип построения статистической радиооптики. Рассмотрены фундаментальные пределы точности, связанные с квантовыми эффектами при регистрации оптических излучений; выполнен анализ особенностей влияния открытого атмосферного канала на работу лазерных систем информационного обмена; исследованы эффекты замираний при распространении лазерных излучений. **Практическая значимость:** формирование и развитие статистической теории лазерных информационных систем позволяет оптимизировать существующие и синтезировать качественно новые алгоритмы радиооптической обработки информации для систем лазерной связи, локации, навигации и управления.

Ключевые слова — обработка информации, статистическая радиооптика, фотоника, лазерные информационные технологии, обобщенные модели, вероятностный анализ, потенциальные характеристики.

Введение

В настоящее время уже нет особой необходимости обосновывать или оправдывать актуальность исследований в области радиооптики. За прошедшие после появления лазеров десятилетия накопилось много убедительных доказательств того, что методы радиофизики и радиотехники эффективно могут использоваться в оптическом диапазоне. В свою очередь, оптические методы активно внедряются в информационные технологии, системы интеллектуального управления, локацию, навигацию, вычислительную технику и технику связи. Разнообразное взаимное сочетание методов радио и оптики, электроники и фотоники дает основу для развития качественно новых лазерных технологий получения, преобразования, передачи и обработки больших массивов информации при высоком быстродействии и при использовании как аналоговых, так и цифровых представлений.

Одновременно с этими особенностями необходимо отметить, что общее развитие радиооптики сопровождается многочисленными парадоксами [1]. Некоторые из этих парадоксов достаточно наглядны, и их полезно выделить.

1. Уже на протяжении почти трех десятилетий в различных монографиях, учебниках и статьях обсуждаются преимущества оптического диапазона частот для передачи, преобразования и обработки информации. Однако многие из этих преимуществ оптики так и остаются до настоящего времени лишь «потенциальными» — нереализованными.

2. Одной из характерных особенностей развития современной науки и техники является

междисциплинарный подход к исследованиям и взаимопроникновение наук. За последние годы в смежных с радиофизикой и оптикой областях сформировались и активно развиваются новые самостоятельные направления исследований: оптическая электроника [2–4] и фотоника [5, 6], оптическая голография [7, 8], оптическая связь [9, 10], оптическая радиометрия [11], акустооптика и электрооптика [12–14], лазерная локация [15–17] и лазерная дальнометрия [18], лазерная спектроскопия [19], лазерная интерферометрия [20, 21], адаптивная оптика [22, 23], компьютерная оптика [24], информационная оптика [25]. Однако формирование многочисленных новых научных направлений до настоящего времени не привело к ожидаемым «революционным» преобразованиям техники.

3. Основным стимулом развития радиооптики всегда являлись информационные задачи — задачи измерений, передачи, приема, управления и обработки информации. Все эти задачи по своей постановке относятся к классу статистических задач. Однако, несмотря на это, статистическая оптика [26, 27], статистическая акустооптика [28], статистическая радиооптика [29, 30] и сегодня находятся лишь на этапе формирования.

4. Подавляющее большинство оптических процессов строится сегодня на основе эвристических принципов. Информационная оптика и лазерные информационные технологии в значительной степени развиваются по направлению исследования новых материалов и новых физических эффектов. В области передачи, управления и обработки информации оптика и фотоника стремятся повторить и улучшить уже существующие

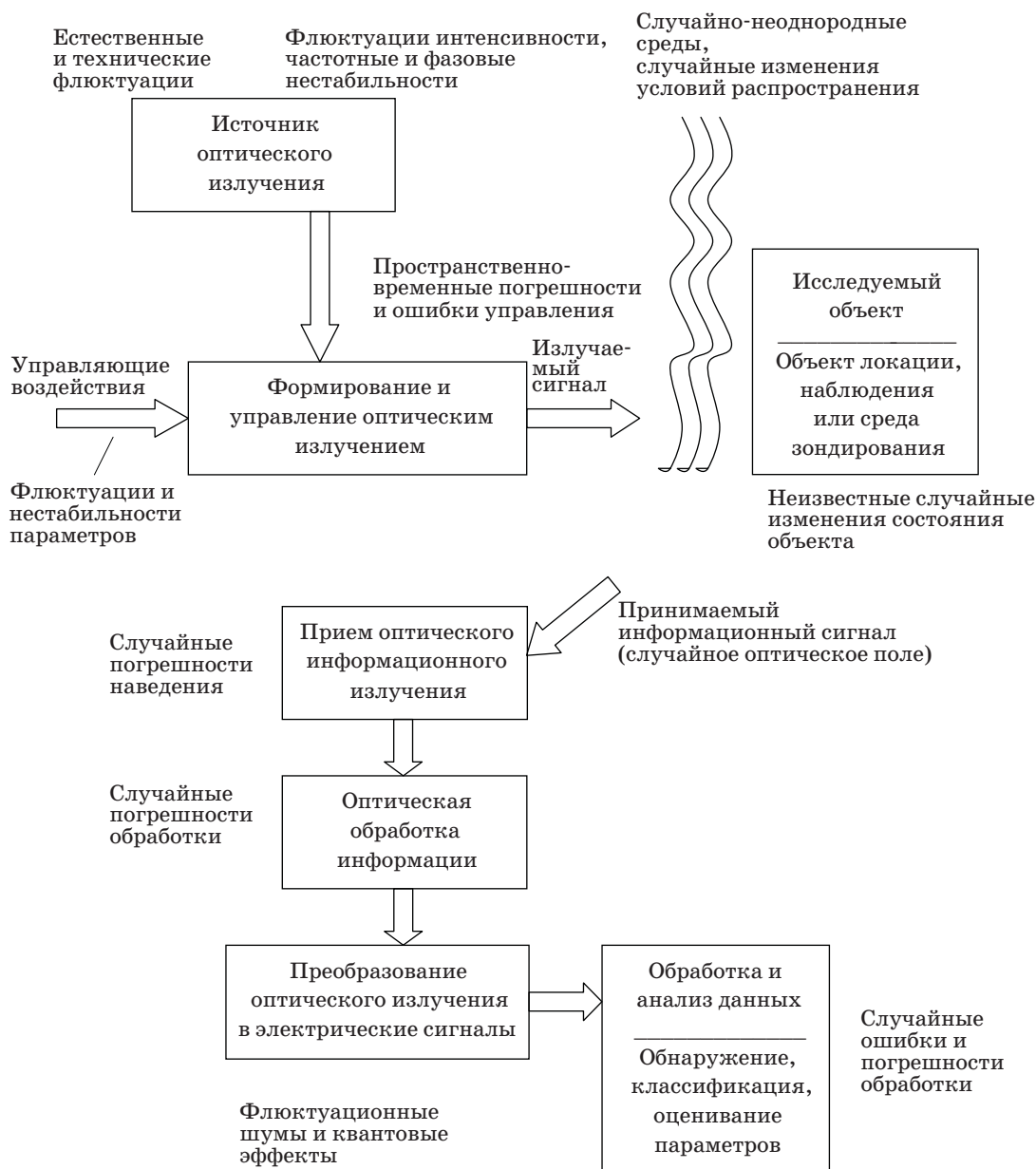
ющие радиоэлектронные системы или улучшить отдельные известные алгоритмы. Однако догнать радиотехнику и электронику эвристическими методами — весьма сомнительный путь. В современной статистической теории существуют хорошо освоенные общие методы оптимального синтеза и анализа информационных и информационно-управляющих систем.

Перечисление подобных противоречий в развитии радиооптики можно продолжать. Все они являются следствием значительного разрыва между физическими основами радиооптики, технологическими проблемами изготовления оптических компонентов и общей статистической теорией обработки информации.

Один из возможных путей устранения перечисленных противоречий — это обобщение многочисленных разрозненных результатов в области радиооптики и построение единой статистической теории лазерных информационных систем и радиооптических методов получения, преобразования, передачи и обработки информации.

Статистические явления в лазерных информационных системах

Построение любой радиооптической системы определяется содержанием и условиями решаемой задачи, видом информационных процессов, требуемыми алгоритмами обработки и анализа



■ Рис. 1. Обобщенная модель радиооптической системы получения, преобразования и обработки информации

данных. Большое разнообразие практических задач неизбежно приводит и к большому разнообразию существующих систем. Однако несмотря на это основной принцип построения радиооптической обработки информации условно можно представить в виде некоторой обобщенной модели. Рассмотрение таких моделей позволяет выделить характерные особенности и возможные подходы к общему анализу различных по своему назначению лазерных систем.

Независимо от конкретных алгоритмов обработки, при решении задач лазерной локации и лазерной навигации, лазерной связи, передачи и обработки информационных процессов, как правило, радиооптические системы содержат некоторые общие, типовые узлы и преобразования. К ним, в частности, относятся источники оптического излучения, операции формирования заданной пространственно-временной структуры излучения, операции управления оптическим излучением, операции приема, обработки и преобразования оптических полей в электрические сигналы. В целом структуру таких систем можно представить обобщенной моделью (рис. 1). Помимо основных преобразований, на этой модели указаны и некоторые случайные факторы, характерные для радиооптики.

Представленная упрощенная модель достаточно наглядно показывает, что случайных факторов и статистических явлений в радиооптике так же много, как и в других областях, связанных с проблемами передачи, приема и обработки информации. С одной стороны, это объясняется тем, что задачи обработки информации по своему содержанию относятся к классу статистических задач. С другой стороны, многие физические эффекты в радиооптике, связанные с генерацией и регистрацией оптических излучений, распространением излучения, формированием оптических изображений, взаимодействием оптических полей с веществом, носят принципиально статистический характер.

Тепловые источники оптического излучения по своей физической природе относятся к генераторам случайного поля. Квантовые генераторы, источники лазерного излучения в силу квантовых эффектов, при всей их стабильности и монохроматичности излучения, неизбежно обладают флуктуациями интенсивности, частотными и фазовыми нестабильностями.

Ясно, что любая реальная система не может быть идеальной. Оценка естественных флуктуаций, технических и технологических погрешностей, случайных внешних и внутренних воздействий может проводиться только на основе вероятностных и статистических методов. Реальные системы управления и обработки информации, как правило, работают в условиях высокой апри-

орной неопределенности и в сложной меняющейся помеховой обстановке. Без вероятностных моделей и статистических методов здесь трудно синтезировать оптимальные алгоритмы и трудно оценивать результаты обработки информационных процессов.

Особенности построения «статистической радиооптики»

Рассмотрение обобщенной модели показывает, что построение и развитие статистической теории лазерных информационных систем должно проходить по двум основным направлениям — вероятностного анализа и статистического синтеза.

Вероятностный анализ должен включать в себя несколько самостоятельных разделов:

- анализ источников радио- и оптического излучения;
- анализ принципов пространственно-временного управления излучением;
- анализ эффектов распространения лазерных излучений в случайно-неоднородных средах;
- анализ принципов ввода информационных сигналов в оптические системы обработки;
- анализ типовых операций радиооптической обработки;
- анализ операций приема и преобразования оптических излучений в электрические сигналы;
- анализ операций последетекторной обработки сигналов.

Статистический синтез должен включать в себя следующие разделы:

- синтез оптимальной структуры информационных сигналов;
- синтез оптимальных алгоритмов передачи, приема и обработки информации.

Методы статистического синтеза должны позволять находить наилучшую структуру сигналов для достижения поставленной цели и наилучшую структуру радиооптических систем. Типовыми примерами подобных задач является синтез оптимальных сигналов для лазерных систем зондирования, локации, передачи информации или синтез оптимальных алгоритмов обнаружения, различения и оценивания параметров оптических излучений в сложной помеховой обстановке.

В свою очередь, чтобы решение задач вероятностного анализа и статистического синтеза было возможно, необходимо на основе теории случайных функций построить удобные для лазерных информационных технологий вероятностные модели процессов и вероятностные модели систем.

Важно подчеркнуть, что для формирования названных разделов статистической радиооптики накоплено уже много полезных результатов.

Во многих смежных с радиооптикой областях к настоящему времени созданы самостоятельные статистические теории. Наглядными примерами являются здесь статистическая физика [31, 32], статистическая радиофизика [29, 33] и статистическая радиотехника [34–36], статистическая радиолокация [37, 38], статистическая теория связи [39, 40], статистическая и общая теория оптимального управления [41–44]. Созданы хорошие основы для развития статистической теории лазерной связи [45, 46], статистической оптики [27, 30] и статистической акустооптики [28].

Все перечисленные направления создавались трудами многих исследователей на протяжении нескольких десятков лет (в основном после 40-х годов прошлого столетия). Часть полученных в смежных областях результатов может успешно использоваться и в радиооптике. Однако в самой радиооптике много своих отличительных особенностей, которые не рассматривались в смежных статистических теориях и без рассмотрения которых трудно оценить даже потенциальные возможности существующих лазерных систем и оптических процессоров.

Фундаментальные пределы точности в радиооптике

В радиооптике, фотонике и лазерных информационных системах, как и в любой другой области, существуют свои принципиальные ограничения — фундаментальные пределы. Эти пределы связаны с физикой используемых явлений и достаточно наглядно проявляются при анализе операций регистрации оптических полей или операций преобразования оптических излучений в информационные электрические сигналы.

В основе таких операций обычно лежит эффект фотоэмиссии, который по своей физической природе имеет принципиально вероятностный характер [27, 47]. При его описании оптическое поле представляется в виде совокупности фотонов, среднее число которых характеризует энергию поля. Каждый фотон при взаимодействии с фоточувствительной поверхностью детектора инициирует с некоторой вероятностью вылет электрона. Появление фотоэлектронов — носителей тока — приводит к появлению первичного тока детектора.

Хорошо известно [47], что если в интервале времени $[t_0, t_0 + T]$ на точечный фотодетектор с квантовой эффективностью η воздействует оптическое излучение с постоянной интенсивностью $I = I(t, r_0) = \text{const}$, то число фотоэлектронов n на выходе детектора характеризуется распределением Пуассона:

$$p(n; T) = \frac{1}{n!} (\eta IT)^n \exp(-\eta IT), \quad n = 1, 2, \dots \quad (1)$$

Этот результат показывает, что при фиксированных значениях η , I и T поток фотоотсчетов $n(t)$ является случайным пуассоновским потоком событий, для которого математическое ожидание m_n и дисперсия σ_n^2 соответственно равны

$$m_n = M\{n(t)\} = \eta IT; \quad \sigma_n^2 = D[n(t)] = \eta IT. \quad (2)$$

В реальных (неидеализированных) ситуациях работа любой радиооптической системы происходит при случайных входных воздействиях. Если при этом интенсивность $I(t, r_0)$ в точке пространства с координатами r_0 описывается некоторой плотностью вероятности $p_I(I)$, то поток фотоотсчетов $n(t)$ на выходе детектора будет характеризоваться распределением

$$p(n; T) = \frac{1}{n!} \int_0^\infty (\eta IT)^n \exp(-\eta IT) p_I(I) dI. \quad (3)$$

Это выражение соответствует известной формуле Мандела [48]; в статистической оптике [27, 29, 47] она лежит в основе большинства исследований по анализу процессов фотоэлектрического преобразования.

Вместе с тем плотность вероятности (3) по существу описывает здесь дважды стохастический пуассоновский поток (модель процесса Д. Кокса [49, 50]), и следовательно, независимо от конкретного вида функции $p_I(I)$ можно сразу же сделать некоторые общие выводы относительно статистики потока фотоотсчетов $n(t)$.

Так, в частности, при любых распределениях $p_I(I)$ входной интенсивности $I(t, r_0)$ математическое ожидание и дисперсия числа фотоотсчетов определяются как

$$m_n = M\{n(t)\} = \eta T m_I; \quad \sigma_n^2 = M\{n^2(t)\} - M^2\{n(t)\} = \eta T m_I + \eta^2 T^2 \sigma_I^2, \quad (4)$$

где m_I и σ_I^2 — параметры процесса $I(t, r_0)$.

Сравнение результатов (4) и (2) позволяет выделить одно из наиболее важных свойств потока фотоотсчетов: случайные изменения интенсивности $I(t, r_0)$ входного оптического излучения всегда приводят к увеличению дисперсии выходного потока фотоэлектронов. Дисперсия σ_n^2 числа фотоотсчетов $n(t)$ состоит из двух различных по физическому смыслу слагаемых. Первое из этих слагаемых $\eta T m_I$ связано с квантовыми флуктуациями при детектировании, а второе слагаемое $(\eta T)^2 \sigma_I^2$ пропорционально дисперсии σ_I^2 интенсивности $I(t, r_0)$ входного излучения и, следовательно, связано с волновыми характеристиками оптического поля.

Если в качестве меры рассеяния для потока фотоотсчетов ввести коэффициент вариации

$$V_n = \frac{\sigma_n}{m_n} = \frac{D^{1/2}[n(t)]}{M\{n(t)\}}, \quad m_n > 0, \quad (5)$$

то на основании результатов (3) и (4) при случайных интенсивностях $I(t, r_0)$ для V_n будет справедливо общее выражение

$$V_n = (\eta T m_I)^{-1} \sqrt{\eta T m_I + (\eta T)^2 \sigma_I^2}. \quad (6)$$

Из него видно, что если флуктуации оптического поля малы и $\sigma_I \ll m_I$, то коэффициент V_n стремится к своей нижней границе:

$$V_n = V_{n \min} = (\eta T m_I)^{-1/2}, \quad (7)$$

а распределение фотоотсчетов (3) приближается к простому пуассоновскому распределению (1) с параметром $m_n = \sigma_n^2 = \eta T I$. Флуктуации потока $n(t)$ обусловлены при этом в основном эффектами фотоэмиссии. В случае «сильных» оптических полей, когда флуктуации интенсивности $I(t, r_0)$ возрастают и $\sigma_I > m_I$, а значение $(\eta T)^2 \sigma_I^2 \gg \eta T m_I$, для коэффициента вариации (6) можно записать

$$V_n = (\eta T m_I)^{-1} \sqrt{\eta T m_I + (\eta T)^2 \sigma_I^2} \approx \sigma_I m_I^{-1} = V_I. \quad (8)$$

Физически ясно, что в подобном случае рассеяние потока фотоотсчетов $n(t)$ определяется коэффициентом вариации V_I интенсивности излучения $I(t, r_0)$, т. е. случайные изменения интенсивности $I(t, r)$ вносят здесь основной вклад во флуктуации числа фотоотсчетов. Важно подчеркнуть, что при выполнении условия (8) дискретное распределение $p(n; T)$ будет приближаться к непрерывному распределению $p_I(I)$ интенсивности $I(t, r_0)$ или, более точно, к распределению процесса $I^*(t, r_0) = \eta T I(t, r_0)$. Квантовые эффекты фотоэлектрического преобразования здесь играют незначительную роль, и поток фотоотсчетов может рассматриваться как непрерывный случайный процесс $I^*(t, r_0)$.

Конкретизируем теперь вид оптического излучения и рассмотрим практически важную для задач обработки информации модель типа «сигнал плюс шум». Будем считать, что сигнал $s(t, r)$ представляет собой высокостабильное лазерное излучение, а шум $\xi(t, r)$ — тепловое гауссово поле. Для аддитивной смеси сигнала и шума плотность вероятности $p_I(I)$ интенсивности $I(t, r_0)$ в такой модели удобно представить в виде [28, 51]

$$p_I(I) = \frac{1}{\sigma_\xi^2} \exp\left(-\frac{a^2}{2}\right) \exp\left(-\frac{I}{\sigma_\xi^2}\right) I_0\left(a \sqrt{2 \frac{I}{\sigma_\xi^2}}\right), \quad (9)$$

где $a = A_s / \sigma_\xi$ — отношение сигнал/шум; $I_0(x)$ — функция Бесселя нулевого порядка от мнимого аргумента.

Подставив выражение (9) в формулу (3) и выполнив интегрирование, получим окончательно-

ный вариант для статистики потока фотоотсчетов

$$p(n; t) = \frac{k^n}{(1+k)^{n+1}} \exp\left(-a^2 \frac{k}{1+k}\right) L_n\left(-\frac{a^2}{1+k}\right), \quad (10)$$

$$k = \eta T \sigma_\xi^2,$$

где $L_n(y)$ — полиномы Лагерра [29, 34].

Распределение (10) позволяет для рассматриваемой модели $s(t, r_0) + \xi(t, r_0)$ определить основные параметры числа фотоотсчетов

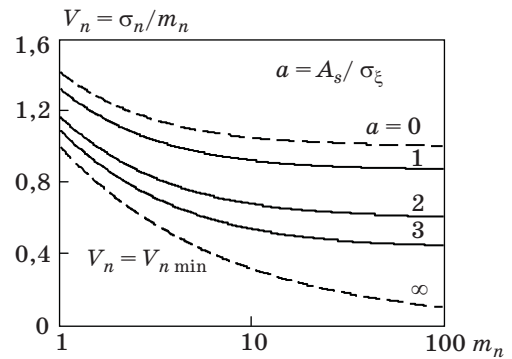
$$m_n = \eta T m_I = \eta T [\sigma_\xi^2 (1+a^2)];$$

$$\sigma_n^2 = m_n \left[1 + m_n \frac{1+2a^2}{(1+a^2)^2}\right];$$

$$V_n = \left[\frac{1}{m_n} + \frac{1+2a^2}{(1+a^2)^2}\right]^{1/2}. \quad (11)$$

В качестве иллюстрации на рис. 2 показан характер изменений коэффициента V_n в зависимости от величины m_n и параметра $a = A_s / \sigma_\xi$.

Из самого определения (5) коэффициента вариации V_n нетрудно заметить, что его величина непосредственно влияет на точность измерений параметров фотоотсчетов $n(t)$, а следовательно, и на точность измерения характеристик интенсивности $I(t, r)$ оптических излучений. В значительной степени именно этим определяется и практическая полезность приведенных в данном разделе результатов. Из формул (11) и рис. 2, в частности, видно, что при возрастании отношения сигнал/шум ($a \gg 1$) значение V_n приближается к своей нижней границе $V_n = V_{n \min}$, а при уменьшении этого отношения ($a \ll 1$) — возрастает до $V_n = V_{n \max}$. Тепловое гауссово излучение обладает наибольшей энтропией, и поэтому значение $V_{n \max}$ соответствует здесь верхней границе для V_n . Иначе говоря, кривые $V_{n \min}$ и $V_{n \max}$



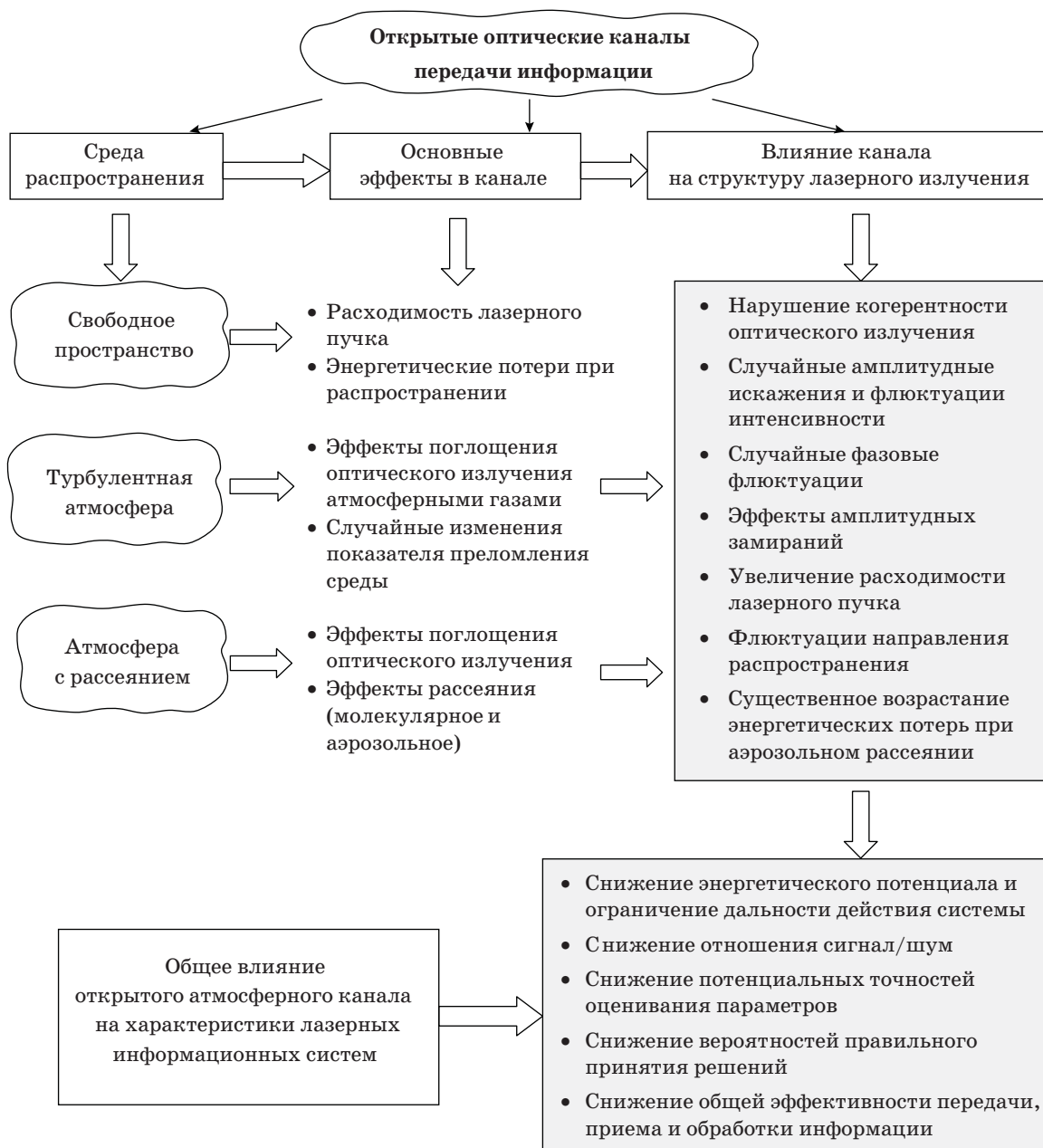
■ Рис. 2. Зависимость коэффициента вариации V_n для числа фотоотсчетов $n(t)$ от соотношения сигнал/шум

задают здесь область допустимых значений для коэффициента вариации V_n и определяют своеобразные «границы точности» для процесса фотоэлектрической регистрации оптических излучений в режиме счета фотонов.

Влияние открытых атмосферных каналов на структуру лазерного излучения

Среди существующего разнообразия лазерных информационных систем целесообразно выделить самостоятельный класс систем с открытым ненаправленным каналом передачи информации.

Средой распространения лазерного излучения в таких каналах может быть свободное космическое пространство (космические каналы), атмосфера (атмосферные каналы), водная среда (подводные каналы). Необходимость исследования открытых каналов наиболее остро проявляется при создании систем лазерной локации и навигации, систем лазерного наведения, прицеливания и целеуказания, систем мобильной лазерной связи, систем лазерного зондирования атмосферы и тропосферы, систем лазерной диагностики различных неоднородных движущихся сред в аэро- и гидродинамике.



■ Рис. 3. Особенности влияния открытых атмосферных каналов на распространение лазерных излучений

Для оптического диапазона открытые каналы передачи представляют собой каналы распространения со случайно-неоднородными средами [52]. Условия распространения лазерного излучения в таких средах оказывают существенное влияние на потенциально достижимые характеристики лазерных систем, предельную дальность действия, эффективность передачи, приема и обработки информации.

Распространение лазерного излучения в условиях турбулентной атмосферы сопровождается, помимо энергетических потерь, эффектами случайного перераспределения энергии и появлением дополнительных амплитудных и фазовых флуктуаций. Такие флуктуации порождаются случайными пространственно-временными изменениями показателя преломления среды вдоль трассы распространения лазерного пучка [53].

Наиболее сложное и многофакторное воздействие атмосферы на распространение лазерных излучений наблюдается в открытых каналах при появлении в атмосфере частиц аэрозоля (облака, туманы, дымки, различные осадки, пыль, загрязнения, ...). Среда распространения становится при этом рассеивающей, и анализ оптических каналов существенно усложняется [52–54]. В каналах с рассеянием значительную роль начинают играть эффекты расходимости лазерного пучка, эффекты поглощения оптического излучения атмосферными газами, эффекты фазовых флуктуаций и амплитудных замираний, эффекты молекулярного и аэрозольного рассеяния. При построении лазерных информационных систем особое влияние оказывает явление аэрозольного рассеяния. Именно этим явлением в значительной степени ограничиваются предельные возможности открытых оптических каналов и именно эффектами аэрозольного рассеяния в значительной степени определяются энергетические потери в лазерных системах.

В целях систематизации и обобщения результатов на рис. 3 представлена структура основных эффектов, влияющих на распространение лазерных излучений в открытых атмосферных

каналах. Здесь выделены наиболее важные для построения физико-математических моделей и вероятностного анализа свойства атмосферного канала, которыми ограничиваются потенциальные возможности лазерных систем связи, управления, локации и навигации.

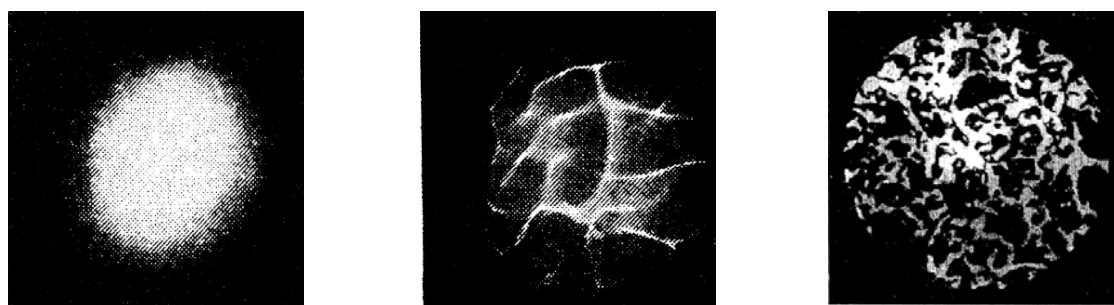
Эффекты случайных замираний в открытых каналах

Многофакторное случайное влияние атмосферных каналов на структуру оптического излучения приводит к снижению общей эффективности передачи, приема и обработки информации. При этом существенно усложняются исследования лазерных систем, и именно поэтому до настоящего времени многие вопросы вероятностного анализа остаются открытыми. Одним из таких вопросов является вопрос анализа эффектов случайных замираний в атмосферных каналах передачи информации.

Случайно-неоднородные среды порождают пространственно-временные искажения структуры лазерного излучения. В качестве иллюстрации на рис. 4 показаны примеры [29, 55] поперечного сечения лазерного луча после его прохождения через турбулентную атмосферу. Светлые и темные участки наглядно отражают здесь пространственную неоднородность интенсивности $I(t, r)$ излучения в некоторой произвольно выбранный момент времени $t = t_1$.

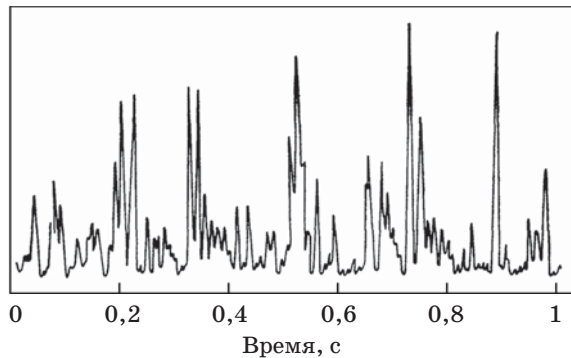
Если прием оптического излучения осуществляется в заданной точке пространства $r = r_0$, то неоднородность среды распространения будет приводить к фазовым и амплитудным флуктуациям оптического поля и соответствующим случайным изменениям временной структуры интенсивности $I(t, r_0)$. На рис. 5, а и б показан характер двух различных реализаций [56, 57] интенсивности $I(t, r_0)$, $t \in [0, T]$ лазерного излучения $I(t, r)$ на выходе канала распространения со случайно-неоднородной средой.

Важной особенностью для информационных систем является здесь то, что пространственно-

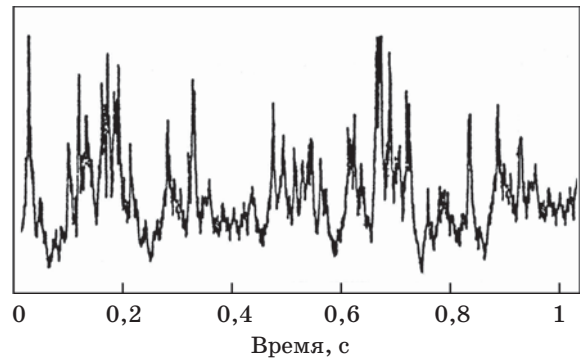


■ Рис. 4. Характер изменения пространственной структуры интенсивности лазерного луча при прохождении турбулентной атмосферы

а) $I(t, r_0)$



б) $I(t, r_0)$



■ **Рис. 5.** Случайные изменения интенсивности лазерного излучения (по данным экспериментальных исследований) при распространении в атмосфере в условиях неоднородной трассы над сушей (а) и на трассе Земля — космос (б)

временная неоднородность структуры оптического излучения порождает эффекты случайных замираний лазерного пучка при его распространении в открытых каналах. Многочисленные теоретические и экспериментальные исследования (например, [53, 56, 57]) позволяют считать, что в качестве основной вероятностной модели для случайных изменений интенсивности $I(t, r)$ может быть выбрана модель логарифмически нормального распределения

$$p_I(I) = (\sigma I \sqrt{2\pi})^{-1} \exp\left(-\frac{\ln^2 I}{2\sigma^2}\right),$$

$$I(t, r_0) \in (0, \infty). \quad (12)$$

В этой модели параметр σ связан с математическим ожиданием m_I и дисперсией σ_I^2 распределения:

$$m_I = M\{I(t, r_0)\} = e^{\sigma^2/2};$$

$$\sigma_I^2 = D[I(t, r_0)] = e^{\sigma^2} (e^{\sigma^2} - 1) = m_I^2 (m_I^2 - 1). \quad (13)$$

При рассмотрении эффектов замираний основной интерес обычно представляют средняя частота замираний и средняя длительность замираний. Эти характеристики зависят от вида распределения интенсивности $I(t, r_0)$, спектрально-корреляционных свойств лазерного излучения и порогового уровня H , относительно которого определяются замирания. Нахождение и вероятностный анализ таких характеристик может быть выполнен на основе общей теории выбросов случайных процессов [51, 58].

Так, в частности, определение среднего числа замираний процесса $I(t, r_0)$ относительно заданного порогового уровня H эквивалентно определению среднего числа отрицательных выбросов $N^-(H, T)$ случайной функции $I(t, r_0)$,

$t \in [t_0, t_0 + T] = [0, T]$ на уровне H . Для вычисления величины $N^-(H, T)$ здесь можно воспользоваться общей формулой [51]

$$N^-(H, T) = - \int_0^T dt \int_{-\infty}^0 I'(t, r_0) p(H, I'; t) dI', \quad (14)$$

в которой функция $p(H, I'; t)$ представляет собой совместную плотность вероятностей $p(I, I'; t) = p(I(t, r_0), I'(t, r_0))$ для значений процесса $I(t, r_0)$ и значений его производной $I'(t, r_0)$ в совпадающие моменты времени при условии $I(t, r_0) = H$.

Такой подход применительно к логарифмически нормальной модели (12), (13) позволяет получить общее выражение для среднего числа замираний $N^-(H, T)$ интенсивности лазерного излучения $I(t, r_0)$ на заданном интервале времени $[t_0, t_0 + T]$ длительностью T при выбранном пороговом уровне H :

$$N^-(H, T) = \frac{T}{2\pi\sigma} \left[\left(\frac{m_I^2 - 1}{m_I^2} \right) (-r_I''(0)) \right]^{1/2} \exp\left(-\frac{\ln^2 H}{2\sigma^2}\right),$$

$$H > 0. \quad (15)$$

$$\text{Значение } -r_I''(0) = -\frac{1}{\sigma_I^2} \frac{d^2}{d\tau^2} R_I(\tau) \Big|_{\tau=0} \text{ определя-$$

ется здесь через корреляционную функцию $R_I(\tau) = \sigma_I^2 r_I(\tau)$ случайного процесса $I(t, r_0)$.

Если воспользоваться теперь результатами (15) и (12), то можно определить среднюю длительность замираний

$$\overline{\tau}^-(H) = \frac{P\{I(t, r_0) \leq H\}}{N^-(H, 1)} =$$

$$= 2\pi [-r_I''(0)]^{-1/2} \Phi\left(\frac{\ln H}{\sigma}\right) \exp\left(\frac{\ln^2 H}{2\sigma^2}\right) \quad (16)$$

и среднюю длительность отсутствия замираний процесса $I(t, r_0)$ на заданном пороговом уровне H

$$\bar{\tau}^+(H) = 2\pi[-r_I''(0)]^{-1/2} \left\{ 1 - \Phi\left(\frac{\ln H}{\sigma}\right) \right\} \exp\left(\frac{\ln^2 H}{2\sigma^2}\right),$$

$$H > 0, \quad (17)$$

где $\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^x \exp(-y^2/2) dy$ — интеграл вероятностей.

Полезно подчеркнуть, что по своему физическому смыслу входящий в формулы (15)–(17) параметр $-r_I''(0)$ связан со спектрально-корреляционными свойствами оптического излучения, и при исследовании замираний его удобно представить в виде [51, 59]

$$-r_I''(0) = -\frac{d^2}{d\tau^2} r_I(\tau) \Big|_{\tau=0} = \gamma^2 \Delta\omega_s^2, \quad (18)$$

где γ — коэффициент формы, а $\Delta\omega_s$ — эффективная ширина спектральной плотности $S(\omega)$ излучения. Если, помимо этого, ввести нормированный пороговый уровень $h = \ln H/\sigma$, то на основе формул (15)–(18) получим простые и удобные для практики результаты:

$$N^-(h, 1) = \frac{1}{2\pi} \gamma \Delta\omega_s \exp(h^2/2), \quad h = \ln H/\sigma;$$

$$\bar{\tau}^+(h) = 2\pi(\gamma \Delta\omega_s)^{-1} \{1 - \Phi(h)\} \exp(h^2/2);$$

$$\bar{\tau}^-(h) = 2\pi(\gamma \Delta\omega_s)^{-1} \Phi(h) \exp(h^2/2). \quad (19)$$

Приведенные результаты (19) наглядно отражают характер зависимости частоты и длительности замираний интенсивности лазерного излучения $I(t, r_0)$ от относительного уровня $h = \ln H/\sigma$, формы спектра (коэффициент γ) и эффективной ширины $\Delta\omega_s$ спектральной плотности $S(\omega)$ излучения.

Рассмотренная модель (12) относится к классу основных вероятностных моделей при описании открытых атмосферных каналов лазерных систем передачи информации. Полученные для этой модели результаты (15)–(19) позволяют достаточно полно исследовать случайную структуру замираний: оценивать вероятность появления замираний, среднюю частоту и среднюю длительность замираний, определять вероятность и среднюю длительность отсутствия замираний на заданном пороговом уровне.

Исследование вероятностной структуры замираний необходимо для анализа свойств и оценки потенциальных возможностей практического использования открытых атмосферных каналов для лазерных информационных систем.

Характеристики замираний существенно влияют на выбор структуры сигналов, выбор методов кодирования и синтез алгоритмов помехоустойчивого приема информации.

Заключение

Безусловно, в данной работе не ставилась цель решить все задачи радиооптики. Здесь лишь предложен один из возможных подходов к построению единой статистической теории лазерных информационных систем и радиооптических методов передачи, приема и обработки информации. Результаты рассмотрения фундаментальных пределов точности и исследование эффектов многофакторного случайного влияния открытых атмосферных каналов на структуру лазерного излучения показывают особенности решения характерных задач вероятностного анализа при построении лазерных информационных систем.

Общее развитие статистической радиооптики позволит:

- исследовать внутренние шумы и флуктуационные эффекты в лазерных системах;
- выполнить анализ работы радиооптических систем в реальной помеховой обстановке;
- определить фундаментальные пределы в радиооптике и оценить потенциальные возможности радиооптических процессоров;
- оптимизировать существующие алгоритмы и синтезировать качественно новые структуры лазерных систем информационного обмена, управления и обработки информации.

Это, в свою очередь, даст возможность рассматривать многие разрозненные задачи лазерной локации, навигации и управления, оптической связи, оптоэлектроники и фотоники с единых позиций — с позиций общей статистической теории обработки информационных процессов.

Работа выполнялась при частичной поддержке Российского фонда фундаментальных исследований в рамках проекта РФФИ № 13-07-12120.

Литература

1. Khimenko V. I. Radio-Optic Signal Processing: Principles of Statistical Theory // Proc. SPIE. 1993. Vol. 2051. P. 711–715.
2. Ярив А. Введение в оптическую электронику / пер. с англ. под ред. О. В. Богданкевича. — М.: Наука, 1983. — 398 с.
3. Хаус Х. Волны и поля в оптоэлектронике / пер. с англ. под ред. К. Ф. Шипилова. — М.: Мир, 1988. — 432 с.
4. Якушенков Ю. Г. Теория и расчет оптико-электронных приборов. — М.: Логос, 1999. — 480 с.

5. Ларкин А. И., Юу Ф. Т. С. Когерентная фотоника. — М.: Бином. Лаборатория знаний, 2007. — 319 с.
6. Салех Б., Тейх М. Оптика и фотоника. Принципы и применения: пер. с англ. — Долгопрудный: Интеллект, 2012. — 760 с.
7. Кольер Р., Беркхарт К., Лин Л. Оптическая голография / пер. с англ. под ред. Ю. И. Островского. — М.: Мир, 1973. — 688 с.
8. Оптическая голография / под ред. Г. Колфилда; пер. с англ. под ред. С. Б. Гуревича. — М.: Мир, 1982. — 736 с.
9. Унгер Г. Г. Оптическая связь / пер. с нем. под ред. Н. А. Семенова. — М.: Связь, 1979. — 264 с.
10. Гауэр Дж. Оптические системы связи / пер. с англ. под ред. А. И. Ларкина. — М.: Радио и связь, 1989. — 504 с.
11. Основы оптической радиометрии / под ред. А. Ф. Котюка. — М.: Физматлит, 2003. — 544 с.
12. Корпел А. Акустооптика: пер. с англ. — М.: Мир, 1993. — 240 с.
13. Балакший В. И., Парыгин В. Н., Чирков Л. Е. Физические основы акустооптики. — М.: Радио и связь, 1985. — 279 с.
14. Ярив А., Юх П. Оптические волны в кристаллах / пер. с англ. под ред. Н. В. Сисакяна. — М.: Мир, 1987. — 616 с.
15. Лазерная локация / под ред. Н. Д. Устинова. — М.: Машиностроение, 1984. — 272 с.
16. Протопопов В. В., Устинов Н. Д. Лазерные локационные системы. — М.: Воениздат, 1987. — 175 с.
17. Основы импульсной лазерной локации / под ред. В. Н. Рождествина. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2006. — 512 с.
18. Лазерная дальнометрия / под ред. В. П. Васильева. — М.: Радио и связь, 1995. — 256 с.
19. Демтредер В. Лазерная спектроскопия / пер. с англ. под ред. И. И. Собельмана. — М.: Наука, 1985. — 608 с.
20. Коронкевич В. П., Соболев В. С., Дубнищев Ю. Н. Лазерная интерферометрия. — Новосибирск: Наука, 1983. — 214 с.
21. Джоунс Р., Уайкс К. Голографическая и спекл-интерферометрия / пер. с англ. под ред. Г. В. Скроцкого. — М.: Мир, 1986. — 328 с.
22. Воронцов М. А., Шмальгаузен В. И. Принципы адаптивной оптики. — М.: Наука, 1985. — 336 с.
23. Адаптивная оптика: сб. ст. / под ред. Э. А. Витриченко. — М.: Мир, 1980. — 456 с.
24. Методы компьютерной оптики / под ред. В. А. Сойфера. — М.: Физматлит, 2003. — 688 с.
25. Информационная оптика / под ред. Н. Н. Евтихьева. — М.: Изд-во МЭИ, 2000. — 612 с.
26. О. Нейл Э. Введение в статистическую оптику / пер. с англ. под ред. П. Ф. Паршина. — М.: Мир, 1966. — 254 с.
27. Гудмен Дж. Статистическая оптика / пер. с англ. под ред. Г. В. Скроцкого. — М.: Мир, 1988. — 528 с.
28. Хименко В. И., Тигин Д. В. Статистическая акустооптика и обработка сигналов. — СПб.: Изд-во СПбГУ, 1996. — 292 с.
29. Ахманов С. А., Дьяков Ю. Е., Чиркин А. С. Введение в статистическую радиофизику и оптику. — М.: Наука, 1981. — 640 с.
30. Ахманов С. А., Дьяков Ю. Е., Чиркин А. С. Статистическая радиофизика и оптика. — М.: Физматлит, 2010. — 428 с.
31. Рейф Ф. Статистическая физика / пер. с англ. под ред. А. О. Вайсенберга. — М.: Наука, 1986. — 336 с.
32. Климонтович Ю. Л. Статистическая физика. — М.: Наука, 1995. — 608 с.
33. Рытов С. М. Введение в статистическую радиофизику. — М.: Наука, 1976. — 496 с.
34. Тихонов В. И. Статистическая радиотехника. — М.: Радио и связь, 1982. — 624 с.
35. Тихонов В. И., Харисов В. Н. Статистический анализ и синтез радиотехнических систем. — М.: Радио и связь, 2004. — 608 с.
36. Тихонов В. И. Оптимальный прием сигналов. — М.: Радио и связь, 1983. — 320 с.
37. Теоретические основы радиолокации / под ред. Я. Д. Ширмана. — М.: Сов. радио, 1970. — 560 с.
38. Сосулин Ю. Г. Теоретические основы радиолокации и радионавигации. — М.: Радио и связь, 1992. — 304 с.
39. Статистическая теория связи и ее практические приложения / под ред. Б. Р. Левина. — М.: Связь, 1979. — 288 с.
40. Информационные технологии в радиотехнических системах / под ред. И. Б. Федорова. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2004. — 768 с.
41. Острем К. Введение в стохастическую теорию управления / пер. с англ. под ред. Н. С. Райбмана. — М.: Мир, 1973. — 322 с.
42. Справочник по теории автоматического управления / под ред. А. А. Красовского. — М.: Наука, 1987. — 712 с.
43. Бесекерский В. А., Попов Е. П. Теория систем автоматического регулирования. — М.: Наука, 1975. — 768 с.
44. Бесекерский В. А. Цифровые автоматические системы. — М.: Наука, 1976. — 576 с.
45. Шереметьев А. Г. Статистическая теория лазерной связи. — М.: Связь, 1971. — 264 с.
46. Гальярди М., Карп Ш. Оптическая связь / пер. с англ. под ред. А. Г. Шереметьева. — М.: Связь, 1978. — 424 с.
47. Клаудер Дж., Сударшан Э. Основы квантовой оптики / пер. с англ. под ред. С. Ахманова. — М.: Мир, 1970. — 428 с.
48. Мандель Д., Вольф Э. Оптическая когерентность и квантовая оптика / пер. с англ. под ред. В. В. Самарцева. — М.: Физматлит, 2000. — 896 с.
49. Кокс Д., Льюис П. Статистический анализ последовательностей событий / пер. с англ. под ред. Н. П. Бусленко. — М.: Мир, 1969. — 312 с.

50. Кингман Дж. Пуассоновские процессы / пер. с англ. под ред. А. М. Вершика. — М.: МЦНМО, 2007. — 136 с.
51. Тихонов В. И., Хименко В. И. Выбросы траекторий случайных процессов. — М.: Наука, 1987. — 304 с.
52. Исимару А. Распространения и рассеяние волн в случайно-неоднородных средах: пер. с англ. — М.: Мир, 1981. Т. 1. — 280 с.; Т. 2. — 318 с.
53. Зуев В. Е., Банах В. А., Покасов В. В. Оптика турбулентной атмосферы. — Л.: Гидрометеиздат, 1988. — 270 с.
54. Тимофеев Ю. М., Васильев А. В. Теоретические основы атмосферной оптики. — СПб.: Наука, 2003. — 474 с.
55. Пратт В. К. Лазерные системы связи / пер. с англ. под ред. А. Г. Шереметьева. — М.: Связь, 1972. — 232 с.
56. Зуев В. Е. Распространение видимых и инфракрасных волн в атмосфере. — М.: Сов. радио, 1970. — 496 с.
57. Оптическая связь (тематический выпуск)//ТИИЭР. 1970. Т. 58. № 10. — 372 с.
58. Тихонов В. И., Хименко В. И. Проблема пересечений уровней случайными процессами. Радиофизические приложения // Радиотехника и электроника. 1998. Т. 43. № 5. С. 501–523.
59. Хименко В. И. Временная когерентность и вероятностная структура интенсивности случайных оптических излучений // Информационно-управляющие системы. 2011. № 1(50). С. 2–8.

UDC 519.2:535

doi:10.15217/issn1684-8853.2015.5.43

Laser Information Systems: Developing a Statistical TheoryKhimenko V. I.^a, Dr. Sc., Tech., Professor, vih.13@yandex.ru^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Purpose: The goal of the paper is summarizing and developing the results in formation of statistical radio-optics which is a unified statistical theory of laser information systems and radio-optical methods of receiving, transformation, transmission and processing of information. **Results:** The main problems and the contradictory nature of radio-optics have been discussed. A generalized model is proposed for radio-optical information processing. A general principle of statistical radio-optics is substantiated. The fundamental accuracy limits associated with quantum effects when registering optical radiation are discussed. The impact of an open atmospheric channel on laser information exchange systems is analyzed. **Practical relevance:** The development of the statistical theory of laser information systems allows us to optimize the existing algorithms and to synthesize completely new algorithms of radio-optical information processing for the systems of laser communication, location, navigation and control.

Keywords — Information Processing, Statistical Radio-optics, Photonics, Laser Information Technologies, Generalized Models, Probabilistic Analysis, Potential Characteristics.

References

1. Khimenko V. I. Radio-Optic Signal Processing: Principles of Statistical Theory. *Proc. SPIE*, 1993, vol. 2051, pp. 711–715.
2. Yariv A. *Introduction to Optical Electronics*. Holt, Rinehart and Winston, 1977.
3. Haus H. *Waves and Fields in Optoelectronics*. New Jersey, Prentice-Hall, 1984.
4. Yakushenkov Yu. G. *Teoriya i raschet optiko-elektronnykh priborov* [Theory and Design of Optoelectronic Devices]. Moscow, Logos Publ., 1999. 480 p. (In Russian).
5. Larkin A. I., Yuu F. T. S. *Kogerentnaya fotonika* [Coherent Photonics]. Moscow, Binom. Laboratoriya znaniy Publ., 2007. 319 p. (In Russian).
6. Saleh B. E. A., Teich M. C. *Fundamentals of Photonics*. New York, John Wiley and Sons, 2007.
7. Collier R. C., Burckhardt C. B., Lin L. H. *Optical Holography*. New York, Academic Press, 1971.
8. *Handbook of Optical Holography*. H. J. Caulfield (Ed.). New York, Academic Press, 1979.
9. Unger H.-G. *Optische Nachrichtentechnik*. Berlin, Elitera, 1976.
10. Gowar J. *Optical Communication Systems*. London, PrenticeHall, 1984.
11. *Osnovy opticheskoi radiometrii* [Fundamentals of Optical Radiometry]. A. F. Kotjuk (Ed.). Moscow, Fizmatlit Publ., 2003. 544 p. (In Russian).
12. Korpel A. *Acousto-optics*. New York, M. Dekker, 1988.
13. Balakshij V. I., Parygin V. N., Chirkov L. E. *Fizicheskie osnovy akustooptiki* [Physical Fundamentals of Acoustooptics]. Moscow, Radio i svyaz' Publ., 1985. 279 p. (In Russian).
14. Yariv A., Yeh P. *Optical Waves in Crystals*. New York, John Wiley and Sons, 1984.
15. *Lazernaya lokatsiya* [Laser Location]. N. D. Ustinov (Ed.). Moscow, Mashinostroenie Publ., 1984. 272 p. (In Russian).
16. Protopopov V. V., Ustinov N. D. *Lazernye lokatsionnye sistemy* [Laser Location System]. Moscow, Voenizdat Publ., 1987. 175 p. (In Russian).
17. *Osnovy impul'snoi lazernoi lokatsii* [Fundamentals of Impulse Laser Location]. V. N. Rozhdestvin (Ed.). Moscow, MGTU im. N. E. Bauman Publ., 2006. 512 p. (In Russian).
18. *Lazernaya dal'nometriya* [Laser Ranging]. V. P. Vasil'ev (Ed.). Moscow, Radio i svyaz' Publ., 1995. 256 p. (In Russian).
19. Demtroder W. *Laser Spectroscopy*. New York, Springer-Verlag, 1982.
20. Koronkevich V. P., Sobolev V. S., Dubnitshev Yu. N. *Lazernaya interferometriya* [Laser Interferometry]. Novosibirsk: Nauka Publ., 1983. 214 p. (In Russian).
21. Jones R., Wykes C. *Holographic and Speckle Interferometry*. Cambridge, Cambridge University Press, 1983.
22. Voroncov M. A., Shmal'gauzen V. I. *Printsipy adaptivnoi optiki* [The Principles of Adaptive Optics]. Moscow, Nauka Publ., 1985. 336 p. (In Russian).
23. *Adaptivnaya optika. Sb. statei* [Adaptive Optics. Collection of Articles]. E. A. Vitrichenko (Ed.). Moscow, Mir Publ., 1980. 456 p. (In Russian).
24. *Metody komp'yuternoi optiki* [Computer Optics Methods]. V. A. Sojfer (Ed.). Moscow, Fizmatlit Publ., 2003. 688 p. (In Russian).

25. *Informatsionnaya optika* [Information Optics]. N. N. Evtihiev (Ed.). Moscow, MEI Publ., 2000. 612 p. (In Russian).
26. O'Neill E. L. *Introduction to Statistical Optics*. Massachusetts, Addison-Wesley Publishing Company, 1963.
27. Goodman J. *Statistical Optics*. New York, John Wiley and Sons, 1985.
28. Khimenko V. I., Tigin D. V. *Statisticheskaya akustooptika i obrabotka signalov* [Statistical Acousto-Optics and Signal Processing]. SPbGU Publ., 1996. 292 p. (In Russian).
29. Ahmanov S. A., D'jakov Yu. E., Chirkin A. S. *Vvedenie v statisticheskuyu radiofiziku i optiku* [Introduction to Statistical Radiophysics and Optics]. Moscow, Nauka Publ., 1981. 640 p. (In Russian).
30. Ahmanov S. A., D'jakov Yu. E., Chirkin A. S. *Statisticheskaya radiofizika i optika* [The Statistical Radiophysics and Optics]. Moscow, Fizmatlit Publ., 2010. 428 p. (In Russian).
31. Reif F. *Statistical Physics: Berkeley Physics Course*. McGraw-Hill Book Company, 1967. Vol. 5.
32. Klimontovich U. L. *Statisticheskaya fizika* [Statistical Physics]. Moscow, Nauka Publ., 1995. 608 p. (In Russian).
33. Rytov S. M. *Vvedenie v statisticheskuyu radiofiziku* [Introduction to Statistical Radiophysics]. Moscow, Nauka Publ., 1976. 496 p. (In Russian).
34. Tikhonov V. I. *Statisticheskaya radiotekhnika* [Statistical Radiotechnics]. Moscow, Radio i svyaz' Publ., 1982. 624 p. (In Russian).
35. Tikhonov V. I., Harisov V. N. *Statisticheskii analiz i sintez radiotekhnicheskikh sistem* [Statistical Analysis and Synthesis of Radio Systems]. Moscow, Radio i svyaz' Publ., 2004. 608 p. (In Russian).
36. Tikhonov V. I. *Optimal'nyi priem signalov* [The Optimal Signal Reception]. Moscow, Radio i svyaz' Publ., 1983. 320 p. (In Russian).
37. *Teoreticheskie osnovy radiolokatsii* [Theoretical Foundations of Radiolocation]. Ya. D. Shirman (Ed.). Moscow, Sovetskoe radio Publ., 1970. 560 p. (In Russian).
38. Sosulin Yu. G. *Teoreticheskie osnovy radiolokatsii i radio-navigatsii* [Theoretical Foundations Radiolocation and Radio Navigation]. Moscow, Radio i svyaz' Publ., 1992. 304 p. (In Russian).
39. *Statisticheskaya teoriya svyazi i ee prakticheskie prilozheniya* [Statistical Communication Theory and its Practical Applications]. B. R. Levin (Ed.). Moscow, Svyaz' Publ., 1979. 288 p. (In Russian).
40. *Informatsionnye tekhnologii v radiotekhnicheskikh sistemakh* [Information Technologies in Radio Systems]. I. B. Fedorov (Ed.). Moscow, MGTU im. N. E. Bauman Publ., 2004. 768 p. (In Russian).
41. Astrom K. *Introduction to Stochastic Control Theory*. Academic Press, 1970. 322 p.
42. *Spravochnik po teorii avtomaticheskogo upravleniya* [Handbook of Control Theory]. A. A. Krasovskij (Ed.). Moscow, Nauka Publ., 1987. 712 p. (In Russian).
43. Besekerskiy V. A., Popov E. P. *Teoriya sistem avtomaticheskogo regulirovaniya* [The Theory of Automatic Control Systems]. Moscow, Nauka Publ., 1975. 768 p. (In Russian).
44. Besekerskiy V. A. *Tsifrovye avtomaticheskije sistemy* [Digital Automatic System]. Moscow, Nauka Publ., 1976. 576 p.
45. Sheremet'ev A. G. *Statisticheskaya teoriya lazernoi svyazi* [Statistical Theory of Laser Communication]. Moscow, Svyaz' Publ., 1971. 264 p. (In Russian).
46. Gagliardi R., Karp S. *Optical Communications*. New York, C. Wiley, 1976.
47. Klaude R. J., Sudarshan E. C. G. *Fundamentals of Quantum Optics*. New York, Benjamin, 1968.
48. Mandel L., Wolf E. *Optical Coherence and Quantum Optics*. Cambridge, Cambridge University Press, 1995.
49. Cox D. R., Lewis P. A. W. *The Statistical Analysis of Series of Events*. London, Methuen; New York, John Wiley, 1966.
50. Kingman J. F. C. *Poisson Processes*. Oxford, Clarendon press, 1993.
51. Tikhonov V. I., Khimenko V. I. *Vybrosy traektorii sluchainykh protsessov* [Trajectories Emissions of Stochastic Processes]. Moscow, Nauka Publ., 1987. 304 p. (In Russian).
52. Ishimaru A. *Wave Propagation and Scattering in Random Media*. Wiley-IEEE Press, 1999, vol. 1.
53. Zuev V. E., Banah V. A., Pokasov V. V. *Optika turbulentnoi atmosfery* [Turbulent Atmosphere Optics]. Leningrad, Gidrometeoizdat, 1988. 270 p. (In Russian).
54. Timofeev Yu. M., Vasil'ev A. V. *Teoreticheskie osnovy atmosferno optiki* [Theoretical Foundations of Atmospheric Optics]. Saint-Petersburg, Nauka Publ., 2003. 474 p. (In Russian).
55. Pratt W. K. *Laser Communication Systems*. New York, Wiley, 1969.
56. Zuev V. E. *Rasprostraneniye vidimykh i infrakrasnykh voln v atmosfere* [Distribution of Visible and Infrared Waves in the Atmosphere]. Moscow, Sovetskoe radio Publ., 1970. 496 p. (In Russian).
57. *Opticheskaya svyaz' (tematicheskii vypusk)* [The Optical Communications (Special Issue)]. *TIER*, 1970, vol. 58, no. 10.
58. Tikhonov V. I., Khimenko V. I. The Problem of Levels Intersections of Stochastic Processes. Radiooptical Applications. *Radiotekhnika i elektronika*, 1998, vol. 43, no. 5, pp. 501–523 (In Russian).
59. Khimenko V. I. Time Coherence and Probability Structure of Random Optical Radiation Intensity. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2011, no. 1(50), pp. 2–8 (In Russian).

МЕТОД ОБРАБОТКИ НЕОДНОРОДНОЙ СТАТИСТИЧЕСКОЙ ИНФОРМАЦИИ О ХАРАКТЕРИСТИКАХ ТОЧНОСТИ СИСТЕМЫ УПРАВЛЕНИЯ

А. А. Ардашов^а, канд. техн. наук, старший научный сотрудник

В. Н. Арсеньев^а, доктор техн. наук, профессор

С. Б. Силантьев^а, канд. техн. наук, профессор

^аВоенно-космическая академия им. А. Ф. Можайского, Санкт-Петербург, РФ

Введение: в процессе опытной отработки систем управления перспективных объектов невозможно обеспечить полную идентичность условий испытаний отдельных образцов из-за проведения доработок, изменения граничных условий и т. д. Одним из путей устранения неоднородности информации, получаемой в процессе испытаний, является приведение результатов испытаний отдельных образцов к некоторым заранее заданным условиям. Качество оценок характеристик системы управления, получаемых по объединенной таким образом выборке, существенно зависит от точности используемых операторов приведения. **Цель:** повышение точности определения операторов приведения по сравнению с известными методами решения этой задачи. **Результаты:** предложен новый подход к объединению неоднородных опытных данных, позволивший повысить точность приведения результатов испытаний к единым условиям и качество оценок характеристик системы. В основу определения оператора приведения положены условия полного совпадения математических ожиданий и максимальной близости ковариационных матриц, характеризующих точность системы в различных условиях. Таким образом обеспечен наиболее полный учет ограниченных опытных данных, полученных в процессе опытной отработки системы управления. Приведен пример оценивания характеристик точности системы управления предложенным методом. **Практическая значимость:** применение полученных результатов позволяет повысить точность оценок характеристик системы управления.

Ключевые слова — система управления, характеристики точности, опытные образцы, условия испытаний, неоднородная статистическая информация.

Введение

Заключительным этапом процесса создания системы управления (СУ) любого объекта являются натурные испытания опытных образцов (ОО). В ходе их проведения ведутся доработки СУ, корректируются алгоритмы управления, могут изменяться граничные условия, полезные нагрузки и т. д. [1–8]. Из-за отличий условий испытаний отдельных опытных образцов получаемые данные о характеристиках СУ являются неоднородными [9–13]. Одним из путей устранения неоднородности является приведение результатов испытаний отдельных образцов к некоторым заранее заданным условиям испытаний [2, 8–10]. Качество оценок характеристик СУ, получаемых по объединенной таким образом выборке, существенно зависит от точности используемых операторов приведения. Целью работы является повышение точности определения операторов приведения по сравнению с известными методами решения этой задачи [8–10].

Постановка задачи

Имеется модель СУ в виде системы нелинейных дифференциальных уравнений

$$\frac{d\hat{\mathbf{X}}}{dt} = \mathbf{F}(\hat{\mathbf{X}}, \mathbf{U}, \hat{\boldsymbol{\lambda}}, t), \quad \mathbf{X}(t_0) = \mathbf{X}_0, \quad (1)$$

где значок « $\hat{}$ » используется для отличия случайной величины от детерминированной; $\hat{\mathbf{X}}(t) = \mathbf{X}_n(t) + \Delta\hat{\mathbf{X}}(t) \in \mathbf{R}^n$ — вектор фазовых координат СУ в момент времени t ; $\mathbf{X}_n(t)$ — его номинальное значение; $\Delta\hat{\mathbf{X}}(t)$ — вектор случайных вариаций фазовых координат СУ относительно номинального значения $\mathbf{X}_n(t)$; $\mathbf{F}(\cdot)$ — в общем случае нелинейная вектор-функция; $\mathbf{U} \in \mathbf{R}^q$ — вектор-функция программ управления; $\hat{\boldsymbol{\lambda}} \in \mathbf{R}^m$ — вектор случайных параметров системы; $t \in (t_0, t_k)$; t_0 и t_k — моменты времени начала и окончания функционирования СУ.

Полагается [4], что распределение вектора вариаций фазовых координат СУ $\Delta\hat{\mathbf{X}}(t_k)$ в конечной точке траектории движения является многомерным нормальным $N(\mathbf{M}_{\Delta\hat{\mathbf{X}}}, \mathbf{K}_{\Delta\hat{\mathbf{X}}})$ с математическим ожиданием $\mathbf{M}_{\Delta\hat{\mathbf{X}}}$ и ковариационной матрицей $\mathbf{K}_{\Delta\hat{\mathbf{X}}}$. Величины $\mathbf{M}_{\Delta\hat{\mathbf{X}}}$ и $\mathbf{K}_{\Delta\hat{\mathbf{X}}}$ характеризуют область рассеивания фазовых координат СУ в момент времени t_k и называются точностными характеристиками (характеристиками точности) СУ. Они зависят от условий испытаний ОО, которые могут отличаться граничными значениями, полезными нагрузками, значениями параметров СУ и т. д.

Пусть в некоторых первых условиях испытано i_1 опытных образцов, а во вторых, отличных от первых, i_2 . Результаты испытаний представлены

соответственно множествами $\Delta X_{1j}, j = \overline{1, i_1}$ и $\Delta X_{2j}, j = \overline{1, i_2}$ значений векторов вариаций фазовых координат СУ в конечных точках траекторий движения.

Необходимо получить оценки M_{12} и K_{12} математического ожидания $M_{\Delta X}$ и ковариационной матрицы $K_{\Delta X}$, характеризующие точность СУ в первых условиях, по результатам испытаний ОО в первых и вторых условиях испытаний.

Определение оператора приведения

Полагается, что в первых и вторых условиях испытаний векторы $\Delta \hat{X}_1$ и $\Delta \hat{X}_2$ вариаций фазовых координат СУ в конечных точках траекторий движения связаны линейной зависимостью:

$$\Delta \hat{X}_1 = P_{12} \Delta \hat{X}_2, \quad (2)$$

где P_{12} — матрица размерности $n \times n$, которая в дальнейшем называется оператором приведения результатов испытаний ОО во вторых условиях к первым условиям.

В силу (2) матрица P_{12} должна удовлетворять двум уравнениям:

$$M_{\Delta \hat{X}_1} = P_{12} M_{\Delta \hat{X}_2}; \quad (3)$$

$$K_{\Delta \hat{X}_1} = P_{12} K_{\Delta \hat{X}_2} P_{12}^T. \quad (4)$$

Оценки параметров $M_{\Delta \hat{X}_1}, M_{\Delta \hat{X}_2}, K_{\Delta \hat{X}_1}$ и $K_{\Delta \hat{X}_2}$

могут быть получены путем многократных статистических испытаний модели (1). Поскольку число испытаний может быть сколь угодно большим, то оценки обозначаются так же, как сами параметры.

Задача определения оператора приведения P_{12} сводится к совместному решению линейного векторного уравнения (3) и нелинейного матричного уравнения (4). В качестве показателей точности полученного решения используются величины

$$\delta_M = \frac{\|M_{\Delta \hat{X}_1} - P_{12} M_{\Delta \hat{X}_2}\|}{\|M_{\Delta \hat{X}_1}\|} 100\% \text{ и}$$

$$\delta_K = \frac{\|K_{\Delta \hat{X}_1} - P_{12} K_{\Delta \hat{X}_2} P_{12}^T\|}{\|K_{\Delta \hat{X}_1}\|} 100\%.$$

Можно показать [4], что система уравнений (3), (4) приводится к виду

$$\mu_1 = Q \mu_2; \quad (5)$$

$$Q Q^T = I, \quad (6)$$

где $\mu_i = D_i^{-1/2} S_i^T M_{\Delta \hat{X}_i}$; $Q = D_1^{-1/2} S_1^T P_{12} S_2 D_2^{1/2}$; $S_i D_i S_i^T = K_{\Delta \hat{X}_i}$; S_i и D_i — ортогональная и диагональная матрицы, состоящие из собственных векторов и собственных значений матрицы $K_{\Delta \hat{X}_i}$ соответственно; $i = 1, 2$.

В общем случае точного решения системы уравнений (3), (4) [и соответствующей системы уравнений (5), (6)] может не существовать. Это может иметь место, когда реальная зависимость между векторами $\Delta \hat{X}_1$ и $\Delta \hat{X}_2$ отличается от линейной (2). В связи с этим предлагается следующий метод приближенного решения системы уравнений (3), (4).

Из (5) видно, что существует бесчисленное множество матриц Q , удовлетворяющих этому уравнению.

Пусть Q_0 — любая матрица той же размерности, что и Q . Тогда общее решение уравнения (5) можно найти, минимизируя функционал $J = \text{tr}[(Q_0 - Q) \times (Q_0 - Q)^T]$ при условии $\mu_1 = Q \mu_2$, где $\text{tr}[\cdot]$ — функция определения следа матрицы.

Решение этой задачи, полученное методом неопределенных множителей Лагранжа, имеет вид

$$Q = Q_0 \left(I - \frac{1}{\mu_2^T \mu_2} \mu_2 \mu_2^T \right) + \frac{1}{\mu_2^T \mu_2} \mu_1 \mu_2^T. \quad (7)$$

Легко проверить, что при любой матрице Q_0 выполняется уравнение (5). В частном случае можно взять в качестве Q_0 единичную матрицу I .

Матрица (7) подставляется в уравнение (6). Очевидно, что в общем случае она не обращает это уравнение в тождество. Однако путем очередного изменения элементов матрицы Q_0 с небольшим шагом ($0,001 \div 0,01$) в сторону уменьшения величины δ_K можно получить достаточно точное с точки зрения показателей δ_M и δ_K решение поставленной задачи.

При этом любая полученная таким образом матрица Q обеспечивает строгое выполнение условия (5), а соответствующая ей матрица

$$P_{12} = S_1 D_1^{1/2} Q D_2^{-1/2} S_2^T \quad (8)$$

является решением уравнения (3).

Оценивание характеристик точности СУ

Результаты испытаний ОО во вторых условиях, приведенные к первым условиям испытаний, $P_{12} \Delta X_{2j}, j = \overline{1, i_2}$, и результаты испытаний ОО в первых условиях $\Delta X_{1j}, j = \overline{1, i_1}$, рассматриваются как выборка из одной генеральной совокупности с законом распределения $N(M_{\Delta \hat{X}_1}, K_{\Delta \hat{X}_1})$. Тогда оценки M_{12} и K_{12} математического ожидания $M_{\Delta \hat{X}_1}$ и ковариационной матрицы $K_{\Delta \hat{X}_1}$, ха-

рактически характеризующие точность СУ в первых условиях испытаний и учитывающие результаты испытаний ОО в первых и вторых условиях, определяют-ся по формулам

$$\mathbf{M}_{12} = \frac{1}{i_1 + i_2} \left(\sum_{j=1}^{i_1} \Delta \mathbf{X}_{1j} + \mathbf{P}_{12} \sum_{j=1}^{i_2} \Delta \mathbf{X}_{2j} \right) = \frac{1}{i_1 + i_2} \left(i_1 \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_1} + i_2 \mathbf{P}_{12} \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_2} \right); \quad (9)$$

$$\mathbf{K}_{12} = \frac{1}{i_1 + i_2 - 2} \times \left\{ \sum_{j=1}^{i_1} \left(\Delta \mathbf{X}_{1j} - \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_1} \right) \left(\Delta \mathbf{X}_{1j} - \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_1} \right)^T + \mathbf{P}_{12} \times \left[\sum_{j=1}^{i_2} \left(\Delta \mathbf{X}_{2j} - \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_2} \right) \left(\Delta \mathbf{X}_{2j} - \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_2} \right)^T \right] \mathbf{P}_{12}^T \right\}, \quad (10)$$

$$\text{где } \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_1} = \frac{1}{i_1} \sum_{j=1}^{i_1} \Delta \mathbf{X}_{1j}; \quad \tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_2} = \frac{1}{i_2} \sum_{j=1}^{i_2} \Delta \mathbf{X}_{2j}.$$

Можно показать, что дисперсии элементов оценок $\mathbf{M}_{12}$ и $\mathbf{K}_{12}$ меньше дисперсий соответствующих элементов оценок, полученных только по результатам испытаний ОО в первых условиях.

Пример. В качестве исследуемого объекта рассматривается ракета-носитель (РН), предназначенная для выведения космических аппаратов (КА) на две различные орбиты. Условия пусков РН отличаются координатами точки выведения. Трехмерный вектор $\Delta \hat{\mathbf{X}}$ состоит из отклонений фазовых координат РН от расчетных значений в конечных точках траекторий выведения. Отклонения измеряются в метрах.

Априорные оценки характеристик точности СУ РН и другие исходные данные заимствованы из работы [10]:

$$\mathbf{M}_{\Delta \hat{\mathbf{X}}_1} = \begin{bmatrix} 25 & 6 & 37 \end{bmatrix}^T; \quad \mathbf{M}_{\Delta \hat{\mathbf{X}}_2} = \begin{bmatrix} 24 & 8 & 12 \end{bmatrix}^T;$$

$$\mathbf{K}_{\Delta \hat{\mathbf{X}}_1} = \begin{bmatrix} 814081 & 561 & 1403298 \\ 561 & 1497 & 1031 \\ 1403298 & 1031 & 2428612 \end{bmatrix};$$

$$\mathbf{K}_{\Delta \hat{\mathbf{X}}_2} = \begin{bmatrix} 776354 & 548556 & 547578 \\ 548556 & 388292 & 387339 \\ 547578 & 387339 & 387377 \end{bmatrix}.$$

Соответствующий этим данным оператор приведения, полученный по формуле (8), имеет вид

$$\mathbf{P}_{12} = \begin{bmatrix} 1,22 & 0,23 & -0,51 \\ 0,38 & -0,83 & 0,30 \\ 0,62 & -0,65 & 2,27 \end{bmatrix}.$$

Величины, характеризующие точность данного решения: $\delta_M = 0$, $\delta_K = 0,04$ %. Для сравнения оператор приведения, найденный путем решения уравнения (4) (в соответствии с [8]), дает $\delta_M = 43$ % и $\delta_K = 0$, а в соответствии с [9] — $\delta_M = 18$ % и $\delta_K = 0,06$ %.

Поскольку реальные пуски РН не проводились, то необходимые статистические данные были получены путем испытаний модели (1). В первых условиях испытаний проведено 6 модельных пусков, результаты которых $\Delta \mathbf{X}_{11} = [-1276 \ -56 \ -2290]^T$; $\Delta \mathbf{X}_{12} = [26 \ -24 \ 51]^T$; $\Delta \mathbf{X}_{13} = [1427 \ 36 \ 2456]^T$; $\Delta \mathbf{X}_{14} = [-729 \ 11 \ -1332]^T$; $\Delta \mathbf{X}_{15} = [684 \ 17 \ 1184]^T$; $\Delta \mathbf{X}_{16} = [54 \ 39 \ 189]^T$, а во вторых — 8 пусков с результатами $\Delta \mathbf{X}_{21} = [-1257 \ -933 \ -910]^T$; $\Delta \mathbf{X}_{22} = [32 \ 10 \ 29]^T$; $\Delta \mathbf{X}_{23} = [1388 \ 999 \ 974]^T$; $\Delta \mathbf{X}_{24} = [-730 \ -521 \ -540]^T$; $\Delta \mathbf{X}_{25} = [667 \ 482 \ 471]^T$; $\Delta \mathbf{X}_{26} = [70 \ 88 \ 81]^T$; $\Delta \mathbf{X}_{27} = [-446 \ -308 \ -316]^T$; $\Delta \mathbf{X}_{28} = [429 \ 258 \ 289]^T$.

По этим данным рассчитаны оценки $\tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_1} = [31 \ 4 \ 43]^T$; $\tilde{\mathbf{M}}_{\Delta \hat{\mathbf{X}}_2} = [19 \ 9 \ 9]^T$ математических ожиданий $\mathbf{M}_{\Delta \hat{\mathbf{X}}_1}$ и $\mathbf{M}_{\Delta \hat{\mathbf{X}}_2}$ соответственно.

Тогда оценки $\mathbf{M}_{12}$ и $\mathbf{K}_{12}$ характеристик точности выведения РН в первых условиях пусков, учитывающие результаты пусков РН в первых и вторых условиях, в соответствии с формулами (9), (10) примут вид

$$\mathbf{M}_{12} = \begin{bmatrix} 26 & 3 & 34 \end{bmatrix}^T;$$

$$\mathbf{K}_{12} = \begin{bmatrix} 810956 & 6662 & 1419402 \\ 6662 & 820 & 12178 \\ 1419402 & 12178 & 2486901 \end{bmatrix}.$$

Очевидно, что дисперсии элементов оценок $\mathbf{M}_{12}$ и $\mathbf{K}_{12}$, учитывающих результаты 14 пусков РН в первых и вторых условиях, будут меньше дисперсий соответствующих элементов оценок $\mathbf{M}_{\Delta \hat{\mathbf{X}}_1}$ и $\mathbf{K}_{\Delta \hat{\mathbf{X}}_1}$, полученных только по результатам шести пусков РН в первых условиях.

Заключение

Приведение разнородных опытных данных к единым условиям испытаний позволяет повысить качество оценок характеристик точности СУ. Результат решения этой задачи существенно зависит от ошибок определения оператора приведения. Предложенный в данной статье подход, как показали многочисленные исследования, во всех случаях обеспечивает нулевую ошибку δ_M и, как правило, малое значение ошибки δ_K . Он может быть использован при решении ряда смежных задач. Так, результаты испытаний ОО в первых условиях могут быть пересчитаны на вторые. Или же, в более общем случае, все

испытания в первых и вторых условиях могут быть использованы для повышения качества оценивания характеристик точности СУ в некоторых, отличных от первых двух, третьих условиях испытаний. Возможны и другие, достаточно очевидные постановки задач, связанных с обработкой неоднородной статистической информации.

Литература

1. Поляков А. П. и др. Справочник по эксплуатации космических средств / под ред. А. П. Полякова. 2-е изд., перераб. и доп. — СПб.: ВКА им. А. Ф. Можайского, 2006. — 758 с.
2. Миронов В. И. Эффективность, надежность и испытания систем управления: учеб. пособие / МО СССР. — М., 1981. — 200 с.
3. Элементы теории испытаний и контроля технических систем / под ред. Р. М. Юсупова. — Л.: Энергия, 1978. — 192 с.
4. Эльясберг П. Е. Определение движения по результатам измерений. — М.: Наука, 1976. — 415 с.
5. Буренок В. М., Найденов В. М. Испытательная база: выход из кризиса // Воздушно-космическая оборона. 2009. № 1(44). С. 18–25.
6. Шаракшанэ А. С., Халецкий А. К., Морозов И. А. Оценка характеристик сложных автоматизированных систем. — М.: Машиностроение, 1993. — 272 с.
7. Кринецкий Е. И., Александровская Л. Н., Шаронов А. В., Голубков А. С. Летные испытания ракет и космических аппаратов. — М.: Машиностроение, 1979. — 464 с.
8. Миронов В. И. Задача приведения вариаций фазовых координат нелинейных динамических систем к заданным условиям испытаний // Изв. АН СССР. Техн. кибернетика. 1970. № 3. С. 31–38.
9. Арсеньев В. Н. Определение характеристик точности системы по результатам ее испытаний в различных условиях испытаний // Изв. АН РФ. Техн. кибернетика. 1992. № 2. С. 118–121.
10. Арсеньев В. Н. Оценивание характеристик точности системы управления ракет-носителя по результатам пусков в различных условиях // Изв. вузов. Приборостроение. 2015. № 1. С. 27–32.
11. Половко А. М., Гуров С. В. Основы теории надежности. — СПб.: БХВ-Петербург, 2006. — 704 с.
12. Вентцель Е. С. Теория вероятностей. — М.: Академия, 2003. — 576 с.
13. Пугачев В. С. Теория вероятностей и математическая статистика. — М.: Физматлит, 2002. — 496 с.

UDC 519.2

doi:10.15217/issn1684-8853.2015.5.55

Processing Miscellaneous Statistical Information about Control System Accuracy Characteristics

A. A. Ardashov^a, PhD, Tech., Senior Research Associate, avgust.ar.@yandex.ru

V. N. Arseniev^a, Dr. Sc., Tech., Professor, vladar56@mail.ru

S. B. Silantyev^a, PhD, Tech., Professor, silantev2008@yandex.ru

^aA. F. Mozhaiskii Military Space Academy, 13, Zhdanovskaia St., 197198, Saint-Petersburg, Russian Federation

Introduction: When testing control systems, it is impossible to provide that the test conditions for various samples are fully identical, because of modifications, changes in the boundary conditions, etc. A way to remove the information heterogeneity is reducing the test results for different samples to certain conditions determined in advance. The estimates of the control system characteristics obtained from the selection united in this way significantly depends on the accuracy of the reduction operators. **Purpose:** Increasing the definition accuracy of the reduction operators as compared to the known methods of solving this task. **Results:** A new approach to amalgamate miscellaneous test data has been proposed which helps to increase the accuracy of reducing the test results to unified conditions and improve the quality of estimating the system characteristics. The reduction operator is determined on the base of full coincidence of population means and the maximum proximity of the correlation matrixes characterizing the accuracy of the system under various conditions. This provides the best possible consideration of limited test data obtained in the course of test functioning of the control system. An example of estimating the accuracy characteristics of the control system is given. **Practical relevance:** Applying the results obtained in this paper can help to increase the accuracy of estimating the characteristics of the control system.

Keywords — Control System, Accuracy Characteristics, Prototypes, Test Condition, Miscellaneous Statistical Information.

References

1. Polyakov A. P., et al. *Spravochnik po ekspluatatsii kosmicheskikh sredstv* [The Reference Book on Operation of Space Means]. Saint-Petersburg, VKA im. A. F. Mozhaiskogo Publ., 2006. 758 p. (In Russian).
2. Mironov V. I. *Effektivnost', nadezhnost' i ispytaniia sistem upravleniia* [Efficiency, Reliability and Tests of Control Systems]. MO SSSR Publ., 1981. 200 p. (In Russian).
3. *Elementy teorii ispytanii i kontrolya tekhnicheskikh sistem* [Elements of the Theory of Tests and Control of Technical Systems]. Ed. R. M. Yusupov. Saint-Petersburg, Energiia Publ., 1978. 192 p. (In Russian).
4. Eliasberg P. E. *Opredelenie dvizheniia po rezul'tatam izmerenii* [Definition of the Movement on Results of Measurements]. Moscow, Nauka Publ., 1976. 415 p. (In Russian).
5. Burenok V. M., Naidyonov V. M. Test Base: Recovery from the Crisis. *Vozdushno-kosmicheskaiia oborona*, 2009, no. 1(44), pp. 18–25 (In Russian).
6. Sharakshane A. S., Haletsky A. K., Morozov I. A. *Otsenka kharakteristik slozhnykh avtomatizirovannykh sistem* [Estimation of Characteristics of the Difficult Automated Systems]. Moscow, Mashinostroenie Publ., 1993. 272 p. (In Russian).

7. Krinetsky E. I., Aleksandrovskaia L. N., Sharonov A. V., Golubkov A. S. *Letnye ispytaniia raket i kosmicheskikh apparatov* [Flight Tests of Rockets and Spacecraft]. Moscow, Mashinostroenie Publ., 1979. 464 p. (In Russian).
8. Mironov V. I. The Task of Reduction of Nonlinear Dynamic Systems Phase Coordinates Variations to the Set Conditions of Test. *Izvestiia AN SSSR. Tekhn. kibernetika*, 1970, no. 3, pp. 31–38 (In Russian).
9. Arseniev V. N. *Opredeleniye harakteristik tochnosti systemi po rezul'tatam ee ispitaniy v razlichnykh usloviyakh* [Definition of System Accuracy Characteristics by Results of Its Tests In Various Conditions]. *Izvestiia AN RF. Tekhn. kibernetika*, 1992, no. 2, pp. 118–121 (In Russian).
10. Arseniev V. N. Estimation of Carrier Rocket Control System Accuracy Characteristics by Results of Start-up in Various Conditions. *Izvestiia vuzov. Priborostroenie*, 2015, no. 1, pp. 27–32 (In Russian).
11. Polovko A. M., Gurov S. V. *Osnovy teorii nadezhnosti* [Bases of The Theory of Reliability]. Saint-Petersburg, BKhV-Peterburg Publ., 2006. 704 p. (In Russian).
12. Ventcel E. S. *Teoriia veroiatnostei* [Probability Theory]. Moscow, Akademiia Publ., 2003. 576 p. (In Russian).
13. Pugachev B. C. *Teoriia veroiatnostei i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. Moscow, Fizmatlit Publ., 2002. 496 p. (In Russian).

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

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

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

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

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

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

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

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

Иллюстрации в текст не завершаются и предоставляются отдельными исходными файлами, поддающимися редактированию:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio 4, 5, 2002–2003 (*.vsd); Coreldraw (*.cdr); Excel (*.xls); Word (*.doc); AdobeIllustrator (*.ai); AutoCad (*.dxf); Matlab (*.ps, *.pdf или экспорт в формат *.ai);

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

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

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

В редакцию предоставляются:

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

— экспертное заключение.

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

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

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

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

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

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (<http://i-us.ru/paperrules>) по разным стандартам: Литература — СИБИД РФ, References — один из мировых стандартов.

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

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ

Кому: Редакция журнала «Информационно-управляющие системы»

Тел.: (812) 494-70-02

Эл. почта: i.us.spb@gmail.com

Сайт: www.i-us.ru

БАЗА ЗНАНИЙ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННО-УПРАВЛЯЮЩЕЙ СИСТЕМЫ МНОГОМЕРНЫМ ОБЪЕКТОМ

Д. Ю. Муромцев^а, доктор техн. наук, профессор

А. Н. Грибков^а, канд. техн. наук, доцент

И. А. Куркин^а, ассистент

^аТамбовский государственный технический университет, Тамбов, РФ

Постановка проблемы: разработка и внедрение информационно-управляющих систем многомерными технологическими объектами являются в настоящее время актуальными задачами, поскольку позволяют повысить качество выпускаемой продукции, снизить затраты энергоресурсов, минимизировать материальные затраты, сократить уровень загрязнения окружающей среды и т. д. К наиболее наукоемким этапам разработки информационно-управляющей системы относится создание ее алгоритмического и программного обеспечения, так как применяемый при этом математический аппарат зачастую является очень сложным. Один из эффективных путей решения данной проблемы — использование интеллектуальных баз знаний, обеспечивающих программную реализацию алгоритмов синтеза энергосберегающих управляющих воздействий. **Методы:** разработка структуры базы знаний с использованием фреймовой модели представления знаний. Реализация в базе знаний метода структурного синтеза алгоритма энергосберегающего управления, основанного на совместном применении принципа максимума Понтрягина и метода синтезирующих переменных. **Результаты:** разработана структура фреймовой базы знаний информационно-управляющей системы многомерным объектом. Фреймы базы знаний обеспечивают определение вида функции оптимального управления и получение аналитических зависимостей для расчета ее параметров. Определение вида функции оптимального управления осуществляется с использованием иерархического графа переходов между видами функций оптимального управления с учетом возможных нарушений ограничений на управляющие воздействия. Параметры полученной функции оптимального управления определяются в результате решения системы уравнений, которая составляется в автоматическом режиме из «элементарных» функций, хранящихся в базе знаний. **Практическая значимость:** применение разработанной базы знаний в составе информационно-управляющей системы многомерным объектом обеспечивает решение задачи синтеза энергосберегающих управляющих воздействий с учетом накладываемых на них ограничений.

Ключевые слова — база знаний, информационно-управляющая система, многомерный объект, энергосберегающее управление.

Введение

В настоящее время вопросам разработки информационно-управляющих систем (ИУС) уделяется значительное внимание исследователей, поскольку такие системы находят широкое применение в современной промышленности. Практически все технологические установки, применяемые на производстве, представляют собой сложные многомерные объекты, имеющие многочисленные взаимосвязанные входные и выходные переменные. При решении задач анализа и синтеза оптимального управления (ОУ) многомерными объектами [1] во многих случаях возникают трудности, связанные со сложностью применяемого математического аппарата и значительной трудоемкостью последующей программной реализации разработанного алгоритмического обеспечения. Поэтому при разработке ИУС многомерными технологическими объектами значительное применение находят методы искусственного интеллекта. При этом в состав ИУС, как правило, включается база знаний (БЗ), обеспечивающая реализацию разработанных алгоритмов анализа и синтеза ОУ.

В данной статье рассматриваются вопросы разработки и программной реализации БЗ ин-

теллектуальной ИУС многомерным объектом. Алгоритмическое обеспечение ИУС основано на методе структурного синтеза алгоритма ОУ, который обеспечивает получение аналитических зависимостей для расчета параметров функций ОУ многомерным объектом при наличии ограничений на управляющие воздействия [2].

Алгоритмическое обеспечение ИУС

Математическую постановку задачи оптимального управления (ЗОУ) многомерным объектом можно записать в следующем виде:

$$\dot{\mathbf{z}}(t) = \mathbf{a}\mathbf{z}(t) + \mathbf{b}\mathbf{u}(t); \quad (1)$$

$$\mathbf{z}(t_0) = \mathbf{z}_0 \rightarrow \mathbf{z}(t_k) = \mathbf{z}_k; \quad (2)$$

$$\forall t \in [t_0; t_k] : \mathbf{u}(t) \in [\mathbf{u}^H; \mathbf{u}^B]; \quad (3)$$

$$J = \int_{t_0}^{t_k} f(\mathbf{u}(t), \mathbf{z}(t), t) dt \rightarrow \min. \quad (4)$$

В задаче (1)–(4) применяются следующие обозначения: $\mathbf{a}$, $\mathbf{b}$ — матрицы параметров модели динамики объекта размерности $n_z \times n_z$ и $n_z \times n_u$ соответственно; $\mathbf{z}(t)$ — вектор фазовых координат размерности n_z ; $\mathbf{u}(t)$ — вектор управляющих воз-

действий размерности n_u ; z_0, z_k — векторы начальных и конечных значений фазовых координат размерности n_z ; $[t_0; t_k]$ — временной интервал управления; u^H, u^B — векторы нижних и верхних граничных значений управляющих воздействий размерности n_u ; J — минимизируемый функционал.

Массив исходных данных, необходимый для численного решения задачи (1)–(4), имеет вид

$$R = (a, b, u^H, u^B, z_0, z_k, t_0, t_k). \quad (5)$$

Практическое решение ЗОУ (1)–(4) заключается в том, чтобы для заданного массива исходных данных (5) определить значения компонентов вектора $u(t)$, при которых для модели объекта (1) с учетом (2) и (3) достигается минимальное значение функционала (4).

Для разработки алгоритмического обеспечения ИУС, осуществляющей решение задачи (1)–(4), используется метод структурного синтеза алгоритма ОУ, основанный на совместном применении принципа максимума Понтрягина и метода синтезирующих переменных [3]. Поясним основные этапы данного метода.

Компоненты вектора управляющих воздействий, согласно принципу максимума Понтрягина, могут принимать три возможных значения: 1) нижнее граничное значение u_i^H ; 2) верхнее граничное значение u_i^B ; 3) значение в виде некоторой функции $u_i^f(t)$. При этом возможны 17 различных видов функций ОУ, которые обобщенно можно представить в виде выражения

$$U^s(t) = \begin{cases} f_1^c(t), & t \in [t_0, t_1^H]; \\ f_2^c(t), & t \in [t_1^H, t_2^B]; \\ f_3^c(t), & t \in [t_2^B, t_3^H]; \\ f_4^c(t), & t \in [t_3^H, t_4^B]; \\ f_5^c(t), & t \in [t_4^B, t_k], \end{cases}$$

где $s = 0, 1, \dots, 16$ — номер вида функции ОУ; $f_i^c(t)$ — компоненты видов функций ОУ, значения которых приведены в таблице; $t_1^H, \dots, t_4^B$ — времена переключения.

Следует также заметить, что из рассмотренных видов функций только $U^0(t)$ является видом функции, не имеющим точек переключения.

Метод структурного синтеза алгоритма ОУ обеспечивает автоматизацию получения уравнений синтезирующих переменных в зависимости от вида функции ОУ [4]. Для этого определяется $n_z \cdot 3^{n_u}$ «элементарных» функций (компонентов синтезирующих переменных), из которых ИУС в процессе работы должна составить n_z уравнений для каждой точки переключения.

■ Значения компонентов видов функций ОУ с одним экстремумом

s	$f_1^c(t)$	$f_2^c(t)$	$f_3^c(t)$	$f_4^c(t)$	$f_5^c(t)$
0	$u_i^f(t)$				
1	$u_i^f(t)$	u_i^B			
2	$u_i^f(t)$	u_i^H			
3	u_i^H	$u_i^f(t)$			
4	u_i^B	$u_i^f(t)$			
5	u_i^H	$u_i^f(t)$	u_i^B		
6	u_i^B	$u_i^f(t)$	u_i^H		
7	$u_i^f(t)$	u_i^B	$u_i^f(t)$		
8	$u_i^f(t)$	u_i^H	$u_i^f(t)$		
9	u_i^H	$u_i^f(t)$	u_i^H		
10	u_i^B	$u_i^f(t)$	u_i^B		
11	u_i^H	$u_i^f(t)$	u_i^B	$u_i^f(t)$	
12	u_i^B	$u_i^f(t)$	u_i^H	$u_i^f(t)$	
13	$u_i^f(t)$	u_i^B	$u_i^f(t)$	u_i^H	
14	$u_i^f(t)$	u_i^H	$u_i^f(t)$	u_i^B	
15	u_i^H	$u_i^f(t)$	u_i^B	$u_i^f(t)$	$u_i^f(t)$
16	u_i^B	$u_i^f(t)$	u_i^H	$u_i^f(t)$	$u_i^f(t)$

Данный подход не требует рассмотрения частных значений синтезирующих переменных, количество которых достаточно велико. Например, для одноэкстремальных функций ОУ, даже без учета взаимного отношения времен переключения, оно равно $17 \cdot n_z \cdot n_u$. Таким образом, при решении ЗОУ требуется производить последовательную проверку различных видов функций ОУ на предмет существования решения ЗОУ (1)–(4) с применением данных видов функций. При этом проверяется выполнение ограничений на управляющие воздействия (3). Применительно к рассмотренным выше видам функций ОУ возможны следующие шесть типов нарушений ограничений:

$$u_i(t_k) > u_i^B; u_i(t_k) < u_i^H; u_i(t_0) > u_i^B;$$

$$u_i(t_0) < u_i^H; u_i(t) > u_i^B, t \in (t_0, t_k);$$

$$u_i(t) < u_i^H, t \in (t_0, t_k). \quad (6)$$

Учитывая, что основные затраты времени ИУС расходуются на решение нелинейных уравнений, можно выделить последовательность проверки видов функций ОУ с учетом количества нелиней-

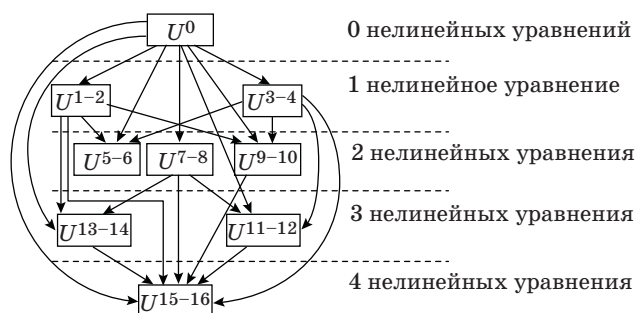


Рис. 1. Иерархический граф переходов между различными видами функций ОУ

ных уравнений. Последовательность проверки можно представить в виде иерархического графа переходов (рис. 1). На каждой итерации выполняется проверка ограничений на управляющие воздействия. При нарушении ограничений, в зависимости от типа нарушений, осуществляется переход от одного вида функции ОУ к другому, более сложному.

Рассмотренный алгоритм синтеза ОУ за счет уменьшения необходимого количества формул позволяет существенно сократить время, затрачиваемое на разработку алгоритмического и программного обеспечения, и снизить вероятность ошибок при кодировании при незначительном увеличении общего времени расчета.

Разработка базы знаний ИУС

Для удобства последующей программной реализации предлагается в качестве модели представления знаний в разрабатываемой БЗ применить фреймовую модель [5], поскольку одним

из основных преимуществ фреймов является способность отражать представление объектных типов данных (классов) в объектно-ориентированных средах разработки программного обеспечения. Следует заметить, что современные понятия класса и объекта, используемые в языках программирования высокого уровня, поддерживающих парадигму объектно-ориентированного программирования, довольно точно соответствуют таким классическим понятиям, как фрейм-образец и фрейм-экземпляр. В связи с этим структуру разрабатываемой БЗ удобно представить в виде диаграммы классов (рис. 2).

База знаний включает шесть классов (фреймов-образцов):

1) FrSourceData — фрейм-образец для ввода и хранения массива исходных данных [A, B, UL, UH, Z0, ZK, T0, TK — соответствующие компоненты массива реквизитов (5); DT — шаг дискретизации; InputData() — функция ввода исходных данных];

2) FrSynthVar — фрейм-образец для расчета значений синтезирующих переменных (NC — количество компонентов синтезирующих переменных; L, CL — массивы для хранения значений синтезирующих переменных и их компонентов («элементарных» функций); CalcL(), CalcCL() — функции для расчета значений L и CL);

3) FrSynthOC — фрейм-образец синтеза оптимальных управляющих воздействий (U — массив для хранения значений управляющих воздействий; CalcParamU(), CalcValU() — функции для расчета параметров и значений ОУ);

4) FrTransitionGraph — фрейм-образец, реализующий алгоритм выбора функции ОУ на основе иерархического графа переходов (NT — массив, содержащий количество переходов по графу для

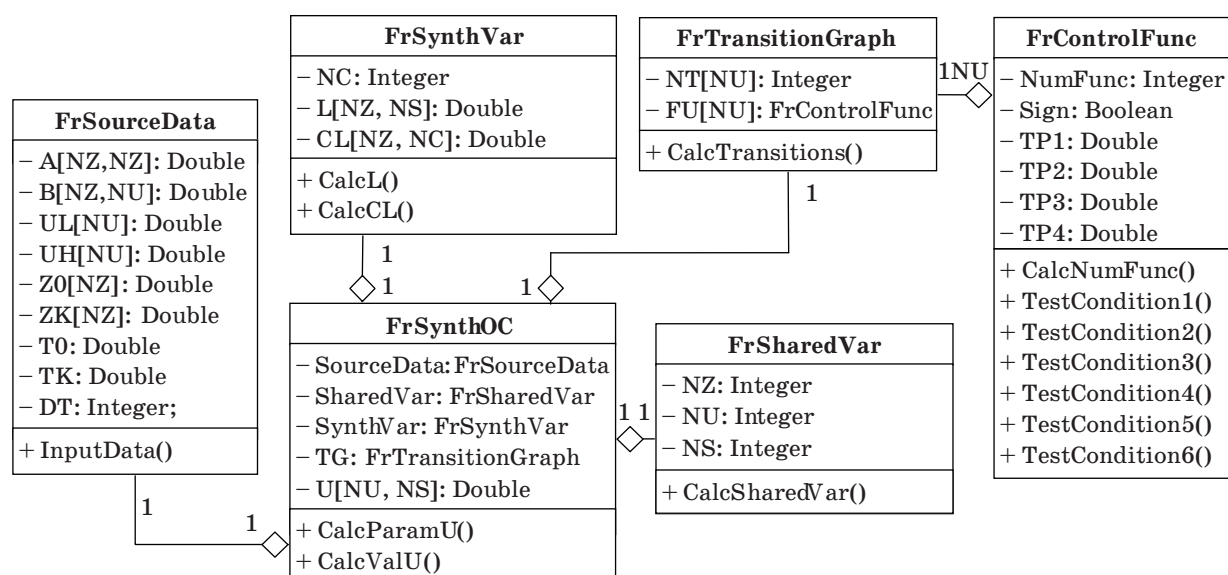


Рис. 2. Структурная схема БЗ ИУС

каждого управляющего воздействия; FU — массив видов функций ОУ; $CalcTransitions()$ — функция расчета переходов по графу);

5) $FrSharedVar$ — фрейм-образец, в котором хранятся значения общих переменных, используемых другими фреймами ($NZ - n_z$; $NU - n_u$; NS — количество шагов на временном интервале управления ($NS = [TK - T0]/DT$));

6) $FrControlFunc$ — фрейм-образец, реализующий определение вида функции ОУ [$NumFunc$ — номер вида функции ОУ; $Sign$ — атрибут, который принимает значение «1» при возрастающей и «0» — при убывающей функции ОУ в начальный момент времени [4]; $TP1, TP2, TP3, TP4$ — значения времен переключения $t_1^H, \dots, t_4^H$; $CalcNumFunc()$ — функция вычисления вида функции ОУ; $TestCondition1(), TestCondition2(), TestCondition3(), TestCondition4(), TestCondition5(), TestCondition6()$ — функции проверки выполнения ограничений (6)].

Алгоритм функционирования ИУС с разработанной БЗ состоит из нескольких этапов. На первом этапе осуществляется ввод исходных данных и расчет значений общих переменных (используются объекты (фреймы-экземпляры) классов $FrSourceData$ и $FrSharedVar$). На втором этапе производится расчет значений синтезирующих переменных и их компонентов («элементарных» функций) с применением объекта класса $FrSynthVar$. На третьем этапе определяется вид

функции ОУ с использованием объектов классов $FrSynthOC$, $FrTransitionGraph$ и $FrControlFunc$. При начальной инициализации объектов за вид функции ОУ принимается $U^0(t)$. Далее выполняется проверка возможных нарушений ограничений (6). При наличии нарушений осуществляются переходы между видами функций ОУ в соответствии с графом переходов (см. рис. 1). На последнем этапе проводится расчет параметров функции ОУ (определение значений времен переключения).

База знаний ИУС программно реализована с использованием среды визуального программирования Delphi 2007 системы CodeGear RAD Studio.

Заключение

В статье рассмотрены особенности построения БЗ ИУС многомерным объектом на базе метода структурного синтеза алгоритма, обеспечивающего получение вида и соотношений для расчета параметров функций ОУ многомерным объектом. Практическая реализация метода в БЗ ИУС обеспечивает более оперативное получение алгоритма синтеза ОУ применительно к многомерным объектам.

Работа выполнена при поддержке гранта Российского фонда фундаментальных исследований, проект № 14-08-00489-а.

Литература

1. Муромцев Ю. Л., Погонин В. А., Гребенников Р. В. Анализ энергосберегающего управления многомерными объектами // Вестник Тамбовского государственного технического университета. 2007. Т. 13. № 4. С. 838–846.
2. Грибков А. Н., Куркин И. А. Аналитический метод получения видов функций и расчета параметров оптимального управления многомерным объектом с учетом ограничений на управляющие воздействия // Информатика и системы управления. 2014. № 3(41). С. 71–83.

3. Муромцев Ю. Л., Ляпин Л. Н., Сатина Е. В. Метод синтезирующих переменных при оптимальном управлении линейными объектами // Изв. вузов. Приборостроение. 1993. № 11-12. С. 19–25.
4. Грибков А. Н., Куркин И. А. Метод структурного синтеза алгоритма расчета параметров функций оптимального управления многомерным объектом // Информатика и системы управления. 2015. № 3(45). С. 72–83.
5. Minsky M. A Framework for Representing Knowledge. — MIT AI Laboratory Memo 306, June, 1974.

UDC 004.896

doi:10.15217/issn1684-8853.2015.5.60

Knowledge Base of an Intelligent Information Control System for a Multidimensional Object

D. Yu. Muromtsev^a, Dr. Sc., Tech., Professor, crems@crems.jesby.tstu.ru

A. N. Gribkov^a, PhD, Tech., Associate Professor, gribkovalexey@yandex.ru

I. A. Kurkin^a, Assistant Professor, allodsli@mail.ru

^aTambov State Technical University, 106, Sovetskaia St., 392000, Tambov, Russian Federation

Purpose: Design and introduction of information control systems for multidimensional technological objects are actual problems because they help to increase the production quality, reduce the energy consumption, minimize the cost, reduce the environmental

pollution, etc. One of the most knowledge-intensive stages in information control system development is creating algorithms and software, as it often involves hard mathematics. A promising way is using intelligent knowledge bases which provide software implementation for the algorithms of synthesizing energy-saving control. **Methods:** The knowledge base structure was developed using the frame model of knowledge representation. To implement the structural synthesis of energy-saving control algorithm in the knowledge base, the Pontryagin maximum principle and the method of synthesizing variables were jointly used. **Results:** A structure has been developed for the frame-based knowledge base of an information control system used in a multidimensional object. The knowledge base frames serve to determine the optimal control function type and to obtain analytical equations for calculating the parameters of this function. The type of the optimal control function is determined using a hierarchical graph of transitions between possible types, taking into account that the constraints on the control actions can be impaired. The parameters of the obtained optimal control function are determined by solving an equation system which is automatically synthesized from "elementary" functions stored in the knowledge base. **Practical relevance:** Using the developed knowledge base as a part of an information control system for a multidimensional object makes it possible to synthesize energy-saving control, taking into account the constraints on the control actions.

Keywords — Knowledge Base, Information Control System, Multidimensional Object, Energy-Saving Control.

References

1. Muromtsev Yu. L., Pogonin V. A., Grebennikov R. V. Analysis of Energy-Saving Control over Multidimensional Objects. *Vestnik Tambovskogo gosudarstvennogo tekhnicheskogo universiteta*, 2007, vol. 13, no. 4, pp. 838–846 (In Russian).
2. Gribkov A. N., Kurkin I. A. Analytical Method for Derivation of Different Functions and Calculation of Optimal Control Parameters for Multidimensional Objects with Limited Control Actions. *Informatika i sistemy upravleniia*, 2014, no. 3(41), pp. 71–83 (In Russian).
3. Muromtsev Iu. L., Liapin L. N., Satina E. V. Method of Synthesizing Variables in the Optimal Control Linear Objects. *Izvestiia vuzov. Priborostroenie*, 1993, no. 11-12, pp. 19–25 (In Russian).
4. Gribkov A. N., Kurkin I. A. Method of Structural Synthesis of Algorithm for Calculating Parameters of Optimal Control Functions for Multidimensional Objects. *Informatika i sistemy upravleniia*, 2015, no. 3(45), pp. 72–83 (In Russian).
5. Minsky M. *A Framework for Representing Knowledge*. MIT AI Laboratory Memo 306, June, 1974.

К статье Н. Н. Васильева и В. С. Дужина «Построение неприводимых представлений симметрической группы $S(n)$ с большими и максимальными размерностями» («Информационно-управляющие системы». 2015. № 3. С. 17–22.).

На странице 21, правый столбец, последний абзац следует читать так: «Работа первого автора выполнена при поддержке грантом РФ 14-11-00581.»

ОПТИМАЛЬНОЕ УПРАВЛЕНИЕ ДВУМЯ ПАРАЛЛЕЛЬНЫМИ FIFO-ОЧЕРЕДЯМИ НА БЕСКОНЕЧНОМ ВРЕМЕНИ

Е. А. Барковский^а, аспирант

А. В. Соколов^{а, б}, доктор физ.-мат. наук, профессор

^аИнститут прикладных математических исследований Карельского научного центра
Российской академии наук, Петрозаводск, РФ

^бПетрозаводский государственный университет, Петрозаводск, РФ

Введение: FIFO-очередь является очень распространенной структурой данных: ее применяют во многих аппаратных и программных приложениях. При разработке различных сетевых устройств и встроенных операционных систем требуется работа с несколькими FIFO-очередями, расположенными в общем пространстве памяти. Также существуют архитектуры многоядерных процессоров, где каждому ядру выделено две FIFO-очереди. Целью исследования является построение и анализ математической модели процесса работы с двумя последовательными циклическими FIFO-очередями в общей памяти, когда на нечетном шаге происходят операции включения элементов в одну из очередей, а на четном шаге — исключения (возможно как последовательное, так и параллельное выполнение операций).

Результаты: сформулирована задача оптимального разбиения общей памяти FIFO-очередей как задача целочисленного программирования, где функция критерия оптимальности задается алгоритмически. Построены математическая и имитационная модели этого процесса для двух очередей и проведены численные эксперименты, основывающиеся на теоретических данных. Математическая модель представлена в виде случайного блуждания по двумерной целочисленной решетке, имеющей отражающие экраны, т. е. мы имеем дело с регулярной однородной марковской цепью. Критерием оптимальности является минимальная средняя доля потерянных при переполнении элементов очередей. Особенностью данного исследования является специфическое выполнение операций над очередями: включение и исключение элементов происходит в зависимости от шага (сделаны поправки для сохранения качеств однородности и регулярности цепи) и выполнение операции возможно параллельно. **Практическая значимость:** с помощью разработанной модели можно найти оптимальное разделение ограниченной общей памяти для повышения стабильности работы системы. Предложенные модели, алгоритмы и разработанный программный комплекс могут применяться при проектировании сетевых устройств, например маршрутизаторов, где потери пакетов являются допустимой, но нежелательной ситуацией. Разделяя общую память для очередей оптимально, мы теряем меньше пакетов, и, как следствие, данные доставляются быстрее.

Ключевые слова — структуры данных, FIFO-очереди, случайные блуждания, регулярные марковские цепи.

Введение

Во многих приложениях, например при разработке различных сетевых устройств и встроенных операционных систем, требуется работа с несколькими FIFO-очередями, расположенными в общем пространстве памяти. Механизм страничной виртуальной памяти здесь не используется, и вся работа происходит в нескольких пулах оперативной памяти. Количество очередей в таких устройствах может достигать нескольких сотен и тысяч, а в будущем, по экспертным оценкам, может достигнуть нескольких миллионов. Для представления FIFO-очередей применяют различные программные или аппаратные решения [1–3].

Отметим также, что среди архитектур многоядерных процессоров есть и такие, где отсутствует кэш-память. Например, в архитектуре AsAP-II каждое ядро имеет два FIFO-буфера, а в архитектуре SEAforth — два стека для хранения данных и адресов возвратов [4]. В этих архитектурах очереди и стеки реализованы циклически и отдельно с возможностью потери элементов при переполнении. Мы же исследуем ситуации, когда для хранения нескольких структур данных исполь-

зуется общая память, что в ряде случаев позволяет снизить потери элементов при переполнении.

Д. Кнут поставил задачу [1] построения и анализа математической модели работы с двумя стеками, растущими навстречу друг другу. Для описания этого процесса были построены модели в виде случайного блуждания в треугольнике [5–10]. В работах [11–14] предлагались модели для последовательного, связанного и страничного способов представления нескольких FIFO-очередей в памяти одного уровня. В этих моделях предполагается, что на каждом шаге дискретного времени с заданными вероятностями происходят некоторые операции со структурами данных. Время выполнения операций — это не случайная величина, а константа, поэтому фиксированным является и шаг времени.

В работе [2] приведены результаты имитационных экспериментов и поставлена задача построить математическую модель процесса работы с несколькими FIFO-очередями в общей памяти, когда операции с очередями выполняются по несколько другому принципу. В данной схеме работы на нечетном шаге допускаются операции включения элементов в одну из n очередей с равными вероятностями, а на четном шаге — опера-

ции исключения элементов из очередей с равными вероятностями. Исключение из пустой очереди не приводит к завершению работы.

Также ставилась задача [2] определить вероятность (как функцию от n и j) того, что очередь, выбранная для операции на j -м шаге, будет пустой, и вычислить математическое ожидание количества элементов в очередях после j операций. В данной задаче не рассматривался конкретный способ представления очередей в памяти, т. е. предполагалось, что очереди могут быть неограниченной длины, что на практике невыполнимо. Эта задача была решена в работе [14].

В данной работе мы предлагаем математическую модель и решаем задачу оптимального разбиения общей памяти ограниченного объема для двух FIFO-очередей в случае последовательного циклического представления очередей. Операции с очередями выполняются по принципу, предложенному выше, и выполнение операций (с заданными вероятностями) может происходить как последовательно, так и параллельно.

В качестве критерия оптимальности рассматривается минимальная доля потерянных элементов при бесконечном времени работы очередей. Эту величину разумно минимизировать, когда переполнение очереди является не аварийной, а стандартной ситуацией (здесь мы подчеркиваем, что в некоторых приложениях при переполнении очереди работа программы заканчивается, и тогда в качестве критерия оптимальности надо рассматривать максимальное среднее время до переполнения памяти). Если очередь занимает всю предоставленную ей память, то все последующие элементы, поступающие в нее, отбрасываются до тех пор, пока не появится свободная память (т. е. пока не произойдет исключение элемента из очереди). Такая схема применяется, например, в работе сетевых маршрутизаторов [3] в том случае, когда по мере увеличения трафика очередь на исходящем интерфейсе маршрутизатора заполняется пакетами. Такое поведение маршрутизатора называется «сбросом хвоста». Потери пакетов приводят к нежелательному результату, поэтому число таких ситуаций необходимо свести к минимуму.

Математическая модель

Пусть в памяти размером в m единиц мы работаем с двумя последовательными циклическими FIFO-очередями с элементами фиксированного размера в одну условную единицу.

Для последовательного представления каждой очереди выделим некоторое количество единиц памяти из общего объема, равного m единиц. Пусть s — количество единиц памяти, выделенных первой очереди, тогда $(m - s)$ — количество единиц памяти, выделенных второй очереди.

Операции, производимые с очередями, выполняются по следующей схеме: на нечетном шаге происходит операция включения элемента в одну из очередей, на четном шаге — операция исключения элемента из какой-либо очереди, причем известны некоторые вероятностные характеристики операций, производимых с очередями. Пусть p_1 и p_2 — вероятности включения элемента в первую и вторую очереди, соответственно p_{12} — вероятность одновременного включения в обе очереди; q_1 и q_2 — вероятности исключения элемента из первой и второй очередей, соответственно q_{12} — вероятность одновременного исключения из обеих очередей.

Поскольку построенная на основе такой постановки задачи марковская цепь [15] не будет регулярной и однородной, два последовательных шага объединяем в один, а также вводим в рассмотрение вероятности выполнения операций, не изменяющих длины очередей. Например, чтение: r_1 — вероятность выполнения операции на нечетном шаге и r_2 — на четном, при этом $r_1 \neq 0, r_2 \neq 0$. Соответственно: $p_1 + p_2 + p_{12} + r_1 = 1, q_1 + q_2 + q_{12} + r_2 = 1$.

Тогда состояния на каждом шаге определяют наступлением одной из следующих комбинаций событий:

- 1) включение в первую, исключение из второй очереди с вероятностью $p_1 q_2$;
- 2) включение во вторую, исключение из первой очереди с вероятностью $p_2 q_1$;
- 3) включение в первую очередь с вероятностью $p_1 r_2 + p_{12} q_2$;
- 4) включение во вторую очередь с вероятностью $p_2 r_1 + p_{12} q_1$;
- 5) включение параллельно в обе очереди с вероятностью $p_{12} r_2$;
- 6) исключение из первой очереди с вероятностью $q_1 r_1 + p_2 q_{12}$;
- 7) исключение из второй очереди с вероятностью $q_2 r_1 + p_1 q_{12}$;
- 8) исключение параллельно из обеих очередей с вероятностью $q_{12} r_1$;
- 9) выполнение над очередями сохраняющих их состояние противоположных операций с вероятностью $r_1 r_2 + p_1 q_1 + p_2 q_2 + p_{12} q_{12}$, при этом сумма вероятностей всех событий равна 1.

Целью исследования является определение оптимального распределения памяти между очередями, когда в качестве критерия оптимальности рассматривается минимальная средняя доля потерянных при переполнении элементов очередей. По закону больших чисел для регулярных цепей Маркова [15] это эквивалентно нахождению решения, доставляющего минимум значению вероятности переполнения памяти на бесконечном промежутке времени.

Обозначим через x и y текущие длины первой и второй очередей соответственно. В качестве

математической модели рассматриваем случайное блуждание в двумерном пространстве по целочисленной решетке в области $0 \leq x \leq s+2$, $0 \leq y \leq m-s+2$ (рис. 1).

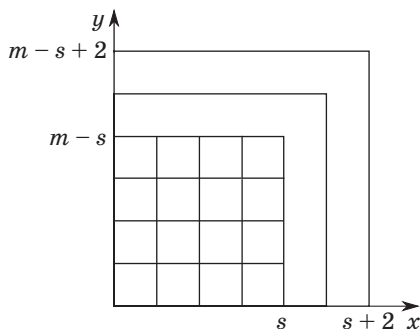
Прямые $x = s+1$, $y = m-s+1$ образуют первый отражающий экран; попадая на эти прямые, мы будем находиться на них до тех пор, пока не произойдет исключение элемента из очереди. Прямые $x = s+2$, $y = m-s+2$ образуют второй отражающий экран, который определен для случаев включения элемента в заполненную очередь и немедленного исключения элемента из этой же очереди. Введением данного экрана учитывается произошедшая потеря элемента, очередь формально переходит на экран, а фактически в область $0 \leq x < s$, $0 \leq y < m-s$, конкретно на прямые $x = s-1$ или $y = m-s-1$.

Определим схему переходов между состояниями. Пусть (x, y) — текущее состояние процесса, тогда блуждание в области $0 \leq x \leq s+1$, $0 \leq y \leq m-s+1$ (включает в себя и первый отражающий экран) можно описать следующим образом:

$$(x, y) \xrightarrow{r_1 r_2} (x', y') = \begin{cases} (x, y), & 0 \leq x \leq s, 0 \leq y \leq m-s \\ (x, y-1), & 0 \leq x \leq s, y = m-s+1 \\ (x-1, y), & x = s+1, 0 \leq y \leq m-s \\ (x-1, y-1), & x = s+1, y = m-s+1 \end{cases};$$

$$(x, y) \xrightarrow{p_1 r_2} (x', y') = \begin{cases} (x, y), & x = s+1, 0 \leq y \leq m-s \\ (x, y-1), & x = s+1, y = m-s+1 \\ (x+1, y), & 0 \leq x \leq s, 0 \leq y \leq m-s \\ (x+1, y-1), & 0 \leq x \leq s, y = m-s+1 \end{cases};$$

$$(x, y) \xrightarrow{p_1 q_1} (x', y') = \begin{cases} (x, y), & 0 \leq x < s, 0 \leq y \leq m-s \\ (x, y-1), & 0 \leq x < s, y = m-s+1 \\ (x+1, y), & x = s+1, 0 \leq y \leq m-s \\ (x+1, y-1), & x = s+1, y = m-s+1 \\ (x+2, y), & x = s, 0 \leq y \leq m-s \\ (x+2, y-1), & x = s, y = m-s+1 \end{cases};$$



■ Рис. 1. Область блуждания

$$(x, y) \xrightarrow{p_1 q_2} (x', y') = \begin{cases} (x, y), & x = s+1, y = 0 \\ (x, y-1), & x = s+1, 0 < y \leq m-s \\ (x, y-2), & x = s+1, y = m-s+1 \\ (x+1, y), & 0 \leq x \leq s, y = 0 \\ (x+1, y-1), & 0 \leq x \leq s, 0 < y \leq m-s \\ (x+1, y-2), & 0 \leq x \leq s, y = m-s+1 \end{cases};$$

$$(x, y) \xrightarrow{p_1 q_{12}} (x', y') = \begin{cases} (x, y), & 0 \leq x < s, y = 0 \\ (x, y-1), & 0 \leq x < s, 0 < y \leq m-s \\ (x, y-2), & 0 \leq x < s, y = m-s+1 \\ (x+1, y), & x = s+1, y = 0 \\ (x+1, y-1), & x = s+1, 0 < y \leq m-s \\ (x+1, y-2), & x = s+1, y = m-s+1 \\ (x+2, y), & x = s, y = 0 \end{cases};$$

$$(x, y) \xrightarrow{p_2 r_2} (x', y') = \begin{cases} (x, y), & 0 \leq x \leq s, y = m-s+1 \\ (x, y+1), & 0 \leq x \leq s, 0 \leq y \leq m-s \\ (x-1, y), & x = s+1, y = m-s+1 \\ (x-1, y+1), & x = s+1, 0 \leq y \leq m-s \end{cases};$$

$$(x, y) \xrightarrow{p_2 q_1} (x', y') = \begin{cases} (x, y), & x = 0, y = m-s+1 \\ (x, y+1), & x = 0, 0 \leq y \leq m-s \\ (x-1, y), & 0 \leq x \leq s, y = m-s+1 \\ (x-1, y+1), & 0 \leq x \leq s, 0 \leq y \leq m-s \\ (x-2, y), & x = s+1, y = m-s+1 \\ (x-2, y+1), & x = s+1, 0 \leq y \leq m-s \end{cases};$$

$$(x, y) \xrightarrow{p_2 q_2} (x', y') = \begin{cases} (x, y), & 0 \leq x \leq s, 0 \leq y < m-s \\ (x, y+1), & 0 \leq x \leq s, y = m-s+1 \\ (x-1, y), & x = s+1, 0 \leq y < m-s \\ (x-1, y+1), & x = s+1, y = m-s+1 \\ (x, y+2), & 0 \leq x \leq s, y = m-s \\ (x-1, y+2), & x = s+1, y = m-s \end{cases};$$

$$(x, y) \xrightarrow{p_2 q_{12}} (x', y') = \begin{cases} (x, y), & x = 0, 0 \leq y < m-s \\ (x, y+1), & x = 0, y = m-s+1 \\ (x, y+2), & x = 0, y = m-s \\ (x-1, y), & 0 < x \leq s, 0 \leq y < m-s \\ (x-1, y+1), & 0 < x \leq s, y = m-s+1 \\ (x-1, y+2), & 0 < x \leq s, y = m-s \\ (x-2, y), & x = s+1, 0 \leq y < m-s \\ (x-2, y+1), & x = s+1, y = m-s+1 \\ (x-2, y+2), & x = s+1, y = m-s \end{cases};$$

$$(x, y) \xrightarrow{p_{12} r_2} (x', y') = \begin{cases} (x, y), & x = s+1, y = m-s+1 \\ (x, y+1), & x = s+1, 0 \leq y \leq m-s \\ (x+1, y), & 0 \leq x \leq s, y = m-s+1 \\ (x+1, y+1), & 0 \leq x \leq s, 0 \leq y \leq m-s \end{cases};$$

$$(x, y) \xrightarrow{p_{12}q_1} (x', y') = \begin{cases} (x, y), & 0 \leq x < s, y = m - s + 1 \\ (x, y + 1), & 0 \leq x < s, 0 \leq y \leq m - s \\ (x + 1, y), & x = s + 1, y = m - s + 1 \\ (x + 1, y + 1), & x = s + 1, 0 \leq y \leq m - s \\ (x + 2, y), & x = s, y = m - s + 1 \\ (x + 2, y + 1), & x = s, 0 \leq y \leq m - s \end{cases};$$

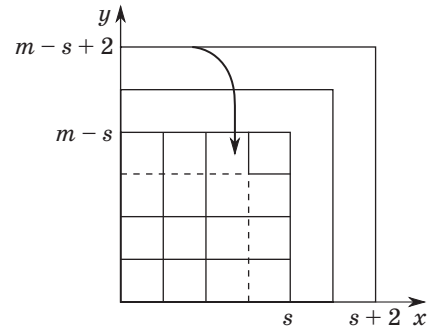
$$(x, y) \xrightarrow{p_{12}q_2} (x', y') = \begin{cases} (x, y), & x = s + 1, 0 \leq y < m - s \\ (x, y + 1), & x = s + 1, y = m - s + 1 \\ (x, y + 2), & x = s + 1, y = m - s \\ (x + 1, y), & 0 \leq x \leq s, 0 \leq y < m - s \\ (x + 1, y + 1), & 0 \leq x \leq s, y = m - s + 1 \\ (x + 1, y + 2), & 0 \leq x \leq s, y = m - s \end{cases};$$

$$(x, y) \xrightarrow{p_{12}q_{12}} (x', y') = \begin{cases} (x, y), & 0 \leq x < s, 0 \leq y < m - s \\ (x, y + 1), & 0 \leq x < s, y = m - s + 1 \\ (x, y + 2), & 0 \leq x < s, y = m - s \\ (x + 1, y), & x = s + 1, 0 \leq y < m - s \\ (x + 1, y + 1), & x = s + 1, y = m - s + 1 \\ (x + 1, y + 2), & x = s + 1, y = m - s \\ (x + 2, y), & x = s, 0 \leq y < m - s \\ (x + 2, y + 1), & x = s, y = m - s + 1 \\ (x + 2, y + 2), & x = s, y = m - s \end{cases};$$

$$(x, y) \xrightarrow{q_1r_1} (x', y') = \begin{cases} (x, y), & x = 0, 0 \leq y \leq m - s \\ (x, y - 1), & x = 0, y = m - s + 1 \\ (x - 1, y), & 0 < x \leq s, 0 \leq y \leq m - s \\ (x - 1, y - 1), & 0 < x \leq s, y = m - s + 1 \\ (x - 2, y), & x = s + 1, 0 \leq y \leq m - s \\ (x - 2, y - 1), & x = s + 1, y = m - s + 1 \end{cases};$$

$$(x, y) \xrightarrow{q_2r_1} (x', y') = \begin{cases} (x, y), & 0 \leq x \leq s, y = 0 \\ (x, y - 1), & 0 \leq x \leq s, 0 < y \leq m - s \\ (x, y - 2), & 0 \leq x \leq s, y = m - s + 1 \\ (x - 1, y), & x = s + 1, y = 0 \\ (x - 1, y - 1), & x = s + 1, 0 < y \leq m - s \\ (x - 1, y - 2), & x = s + 1, y = m - s + 1 \end{cases};$$

$$(x, y) \xrightarrow{q_{12}r_1} (x', y') = \begin{cases} (x, y), & x = 0, y = 0 \\ (x, y - 1), & x = 0, 0 \leq y \leq m - s \\ (x, y - 2), & x = 0, y = m - s + 1 \\ (x - 1, y), & 0 \leq x \leq s, y = 0 \\ (x - 1, y - 1), & 0 \leq x \leq s, 0 \leq y \leq m - s \\ (x - 1, y - 2), & 0 \leq x \leq s, y = m - s + 1 \\ (x - 2, y), & x = s + 1, y = 0 \\ (x - 2, y - 1), & x = s + 1, 0 \leq y \leq m - s \\ (x - 2, y - 2), & x = s + 1, y = m - s + 1 \end{cases};$$



■ Рис. 2. Состояния в области блуждания, соответствующие второму экрану

Переходы со второго отражающего экрана (рис. 2):

1) при $y = m - s + 2$ переходы из состояния (x, y) соответствуют переходам из состояния (x', y') , где $x' = x, y' = y - 3$;

2) при $x = s + 2$ переходы из состояния (x, y) соответствуют переходам из состояния (x', y') , где $x' = x - 3, y' = y$;

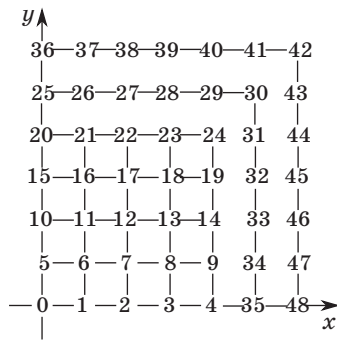
3) при $x = s + 2$ и $y = m - s + 2$ переходы из состояния (x, y) соответствуют переходам из состояния (x', y') , где $x' = x - 3, y' = y - 3$.

Необходимо минимизировать число потерянных элементов при переполнении очередей. Иначе говоря, нужно найти такое s , чтобы математическое ожидание доли времени, которое процесс проводит на отражающих экранах, было минимальным. Таким образом, в случае двух очередей задача оптимального разбиения общей памяти для FIFO-очередей может быть рассмотрена как одномерный случай задачи целочисленного программирования, где функция критерия оптимальности задается алгоритмически. Можно еще дать и такую интерпретацию решаемой задачи. Так как для каждого значения s мы имеем свою цепь Маркова, то можно сказать, что мы решаем задачу нахождения оптимальной цепи Маркова в смысле указанного критерия оптимальности. Для решения данной задачи используются результаты теории регулярных цепей Маркова.

Матрица переходных вероятностей

Представим случайное блуждание в виде регулярной марковской цепи. Пусть P — матрица переходов, $n = (m - s + 3)(s + 3)$ — количество состояний.

Определим нумерацию состояний, как показано на рис. 3. Сначала мы нумеруем состояния в области $x \leq 0 \leq s + 2, y \leq 0 \leq m - s + 2$. Затем состояния, в которых очереди переполняются, — это первый и второй отражающие экраны.



■ Рис. 3. Нумерация состояний при $m = 8, s = 4$

Утверждение. Пронумерованная таким образом матрица P имеет определенную структуру:

$$P = \begin{pmatrix} Q_1 & Q_2 & Q_3 \\ Q_4 & Q_5 & Q_6 \\ & & Q_7 \end{pmatrix},$$

где:

- 1) подматрица Q_1 описывает блуждание в области $x \leq s, y \leq m - s$;
- 2) подматрицы Q_2 и Q_3 описывают переходы из области $x \leq s, y \leq m - s$ на первый и второй отражающие экраны;

Доказательство:

1. База индукции. Пусть общий объем памяти равен $m = 2$; $s = 1$ — количество памяти, выделенное первой очереди. Тогда размер матрицы Q_1 будет 4×4 . Подматрицы A, B, C, D, F имеют размер 2×2 (подматрица C здесь не представлена ввиду маленького значения памяти m). Матрица Q_1 имеет следующую форму:

$$Q_1 = \begin{pmatrix} p_1q_1 + p_2q_2 + p_1q_{12} + & p_2r_2 + p_2q_1 + p_{12}q_1 & p_1r_2 + p_1q_2 + p_{12}q_2 & p_{12}r_2 \\ + p_2q_{12} + p_{12}q_{12} + r_1q_1 + & & & \\ + r_1q_2 + r_1q_{12} + r_1r_2 & & & \\ \hline r_1q_2 + r_1q_{12} + p_1q_{12} & p_1q_1 + r_1q_1 + r_1r_2 & p_1q_2 & p_1r_2 \\ \hline r_1q_1 + r_1q_{12} + p_2q_{12} & p_2q_1 & p_2q_2 + r_1q_1 + r_1r_2 & p_2r_2 \\ \hline r_1q_{12} & r_1q_1 & r_1q_2 & r_1r_2 \end{pmatrix}.$$

2. Индуктивное предположение. Предположим, что для размера памяти m лемма верна. Размерность Q_1 будет $(s + 1)(m - s + 1)$, размерность подматриц — $(s + 1)$.

3. Индуктивный переход. Проверим, что при размере памяти $(m + 1)$ лемма верна. Так как добавилась еще одна единица памяти, то увеличилась одна из очередей. Рассмотрим два случая:

а. Единица памяти попала в первую очередь, размеры памяти m и s увеличились на единицу: $m = m + 1, s = s + 1$. Тогда область случайных блужданий увеличится по оси Ox — будет добавлено $(m - s + 1)$ новых состояний. Размерность матрицы Q_1 увеличится на $(m - s + 1)$, размерность всех подматриц увеличится на единицу, и их количество останется прежним.

3) подматрицы Q_4, Q_5 и Q_6 описывают поведение процесса, когда очереди переполнены;

4) подматрица Q_7 описывает переходы со второго отражающего экрана.

Лемма. Матрица Q_1 размера $(s + 1)(m - s + 1)$ имеет следующий вид:

$$Q_1 = \begin{pmatrix} D & A & O & \dots & O \\ B & C & A & O & \dots \\ O & B & C & \ddots & \dots \\ \vdots & \vdots & \ddots & \ddots & A \\ O & O & \dots & B & F \end{pmatrix},$$

где A, B, C, D, F являются подматрицами размера $(s + 1)$; O — нулевая матрица;

$$A = \begin{pmatrix} p_1r_2 + p_1q_2 + p_{12}q_2 & p_{12}r_2 & 0 & \dots & 0 \\ p_1q_2 & p_1r_2 & p_{12}r_2 & \dots & 0 \\ 0 & p_1q_2 & p_1r_2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & p_{12}r_2 \\ 0 & 0 & \dots & p_1q_2 & p_1r_2 \end{pmatrix}.$$

Мы докажем лемму с помощью математической индукции.

Пусть размерность подматриц будет $(s + 2)$. Проследим за изменением структуры на примере подматрицы A . К ней будут добавлены одна колонка и один столбец, и ее вид будет следующим:

$$A_{(s+2) \times (s+2)} = \begin{pmatrix} A_{(s+1) \times (s+1)} & \vdots \\ \hline 0 & \dots & p_1q_2 & p_1r_2 \end{pmatrix}.$$

Общий вид подматрицы не изменится. Аналогично это можно показать и для остальных подматриц.

б. Единица памяти была добавлена во вторую очередь. Размеры m и s станут $m = m + 1, s = s$.

Область блуждания поднимется вверх по оси Oy , т. е. будет добавлено $(s + 1)$ новых состояний. Размерность матрицы Q_1 увеличится на $(s + 1)$. Размерности подматриц не изменятся, но количество подматриц A , B , C увеличится на один, т. е. общий вид матрицы Q_1 сохранится.

Таким образом, в обоих случаях матрица Q_1 не изменяет свою структуру — лемма доказана.

Используя этот метод, можно доказать, что и остальные подматрицы P сохраняют свой вид. Утверждение доказано.

Некоторые примеры численного анализа

Разработан алгоритм и программа нахождения оптимального разбиения памяти между очередями в зависимости от вероятностных характеристик очередей. Для решения поставленной задачи использовался аппарат управляемых случайных блужданий, регулярных цепей Маркова, система Intel® Math Kernel Library PARDISO. Вычисления производились с помощью кластера КарНЦ РАН.

■ Сравнение потерь

Входные данные	Величина потерь при переполнении ($m = 10$)	
	Оптимальное разбиение	Разбиение пополам ($s = 5$)
$p_1 = 0,25, p_2 = 0,25, p_{12} = 0,25, r_1 = 0,25, q_1 = q_2 = q_{12} = r_2 = 0,25$	0,089 ($s = 5$)	0,089
$p_1 = 0,30, p_2 = 0,20, p_{12} = 0,10, r_1 = 0,40, q_1 = q_2 = q_{12} = r_2 = 0,25$	0,012 ($s = 6$)	0,014
$p_1 = 0,30, p_2 = 0,20, p_{12} = 0,10, r_1 = 0,40, q_1 = q_2 = q_{12} = r_2 = 0,25$	0,019 ($s = 7$)	0,025
$p_1 = 0,30, p_2 = 0,20, p_{12} = 0,10, r_1 = 0,40, q_1 = q_2 = q_{12} = r_2 = 0,25$	0,035 ($s = 7$)	0,046
$p_1 = 0,30, p_2 = 0,20, p_{12} = 0,10, r_1 = 0,40, q_1 = q_2 = q_{12} = r_2 = 0,25$	0,064 ($s = 8$)	0,075

Литература

1. Knuth D. The Art of Computer Programming. — Addison-Wesley Professional, 1997. Vol. 1. — 672 p.
2. Sedgewick R. Algorithms in C++. Parts 1–4. — Addison-Wesley Professional, 1998. — 752 p.
3. Bollapragada V., Murphy C., White R. Inside Cisco IOS Software Architecture. — Cisco Press, 2000. — 240 p.
4. Калачев А. В. Многоядерные процессоры. — М.: БИНОМ, 2014. — 247 с.

Некоторые результаты вычислений приведены в таблице (указанные результаты были подтверждены имитационными экспериментами). Так как аналитическое решение не было получено, нам нужно решать задачу для отдельных m . $m = 10$ используется для примера. Взятые здесь вероятности являются теоретическими — для большей наглядности результатов. На практике же эти вероятности должны быть получены в результате предварительных статистических исследований.

Заключение

Анализируя результаты, можно сказать, что с увеличением вероятности включения в одну очередь (в данном случае — в первую) и уменьшением вероятности включения в другую при разбиении памяти пополам потери при переполнении увеличиваются. При оптимальном разбиении мы выделяем одной очереди s единиц памяти, а второй — $m-s$ единиц памяти. Так, при вероятностях включения 0,35 в одну очередь и 0,15 в другую разница потерь между оптимальным разбиением и разбиением пополам составляет 0,006. То есть из тысячи пакетов мы теряем на 6 пакетов меньше, разбивая память оптимально. А при вероятностях включения 0,45 и 0,05 соответственно — уже на 11 пакетов меньше.

Остается открытым вопрос о том, какой из методов работы с очередями (описанный здесь принцип четных-нечетных шагов или такой, где операции совершаются на любом шаге) будет оптимальным для тех или иных аппаратных или программных решений. В любом случае математические модели таких процессов должны функционировать в дискретном времени. Таким образом, использование классического аппарата теории массового обслуживания (например, основанного на пуассоновском потоке заявок с непрерывным временем) в качестве математических моделей таких систем не представляется оправданным.

Работа выполнена при финансовой поддержке РФФИ, грант 15-01-03404-а.

5. Yao A. C. An Analysis of a Memory Allocation Scheme for Implementing Stacks // SIAM Journal on Computing. 1981. Vol. 10. P. 398–403.
6. Flajolet P. The Evolution of Two Stacks in Bounded Space and Random Walks in a Triangle // Lecture Notes in Computer Science. 1986. Vol. 223. P. 325–340.
7. Louchard G., Schott R. Probabilistic Analysis of Some Distributed Algorithms // Lecture Notes in Computer Science. 1990. Vol. 431. P. 239–253.

8. Louchard G., Schott R., Tolley M., Zimmermann P. Random Walks, Heat Equation and Distributed Algorithms // *Journal of Computational and Applied Mathematics*. 1994. N 53. P. 243–274.
9. Maier R. S. Colliding Stacks: A Large Deviations Analysis // *Random Structures and Algorithms*. 1991. N 2. P. 379–421.
10. Соколов А. В. О распределении памяти для двух стеков // *Автоматизация эксперимента и обработки данных: сб. Петрозаводск: Карельский филиал АН СССР, 1980. С. 65–71.*
11. Aksenova E. A., Sokolov A. V. The Optimal Implementation of Two FIFO-Queues in Single-Level Memory // *Applied Mathematics*. 2011. Vol. 2. P. 1297–1302.
12. Sokolov A. V., Drac A. V. The Linked List Representation of n LIFO-Stacks and/or FIFO-Queues in the Single-Level Memory // *Information Processing Letters*. 2013. Vol. 13. P. 832–835.
13. Аксенова Е. А., Драц А. В., Соколов А. В. Оптимальное управление n FIFO-очередями на бесконечном времени // *Информационно-управляющие системы*. 2009. № 6. С. 46–54.
14. Драц А. В., Соколов А. В. Математический анализ процесса работы с M FIFO-очередями // *Стохастическая оптимизация в информатике*. 2012. Т. 8. № 2. С. 75–82.
15. Kemeny J. G., Snell J. L. *Finite Markov Chains*. — Springer, 1983. — 226 p.

UDC 004.942

doi:10.15217/issn1684-8853.2015.5.65

Optimal Control of Two Parallel FIFO Queues on an Infinite Time

Barkovsky E. A.^a, Post-Graduate Student, barkevgen@gmail.com

Sokolov A. V.^{a, b}, Dr. Sc., Phys.-Math., Professor, avs@krc.karelia.ru

^aInstitute of Applied Mathematical Research of Karelian Research Centre Russian Academy of Sciences, 11, Pushkinskaya St., 185910, Petrozavodsk, Russian Federation

^bPetrozavodsk State University, 33, Lenin St., 185910, Petrozavodsk, Russian Federation

Introduction: FIFO queue is a very common data structure used in many hardware and software applications. In the development of various network devices and embedded operating systems, you often have to work with several FIFO queues located in a shared memory space. Also, there are architectures of multicore processors where two FIFO queues are allocated to one core. The purpose of this research is constructing and analyzing a mathematical model of two serial cyclic FIFO queues in a shared memory with the following protocol: insertion into one of the queues is performed on an odd step, and deletion is performed on an even step (the execution of the operations can be either serial or parallel). **Results:** The problem of optimal partitioning of the shared memory for FIFO-queues is formulated as a problem of integer programming, where the optimality criterion function is defined algorithmically. Mathematical and simulation models of this process were constructed for two queues, and numerical experiments based on theoretical data were performed. The mathematical model was built as a random walk on a two-dimensional integer lattice with reflecting screens, i.e. we deal with a regular uniform Markov chain. The optimality criterion is the minimum average share of lost (due to overflow) elements of the queues. The peculiarity of our approach is a specific execution of the queue operations: insertion and deletion of the elements occur according to the step (some amendments have been made to maintain homogeneity and regularity of the chain), and parallel execution of the operations is also possible. **Practical relevance:** With the help of the developed model, you can find the optimal partition of a limited shared memory to improve the system stability. The proposed models, algorithms and the developed software complex can be used in the design of networking devices, for example, routers, where packet losses are acceptable but unwanted. When we partition a shared memory for queues in the optimal way, we lose fewer packets and, as a result, the data is delivered quicker.

Keywords — Data Structures, FIFO queues, Random Walks, Regular Markov Chains.

References

1. Knuth D. *The Art of Computer Programming*. Addison-Wesley Professional, 1997. Vol. 1. 672 p.
2. Sedgewick R. *Algorithms in C++*. Parts 1–4. Addison-Wesley Professional, 1998. 752 p.
3. Bollapragada V., Murphy C., White R. *Inside Cisco IOS Software Architecture*. Cisco Press, 2000. 240 p.
4. Kalachev A. V. *Mnogoiadernnye protsessory* [Multicore Architectures]. Moscow, BINOM, 2014. 247 p. (In Russian).
5. Yao A. C. An Analysis of a Memory Allocation Scheme for Implementing Stacks. *SIAM Journal on Computing*, 1981, vol. 10, pp. 398–403.
6. Flajolet P. The Evolution of Two Stacks in Bounded Space and Random Walks in a Triangle. *Lecture Notes in Computer Science*, 1986, vol. 223, pp. 325–340.
7. Louchard G., Schott R. Probabilistic Analysis of Some Distributed Algorithms. *Lecture Notes in Computer Science*, 1990, vol. 431, pp. 239–253.
8. Louchard G., Schott R., Tolley M., Zimmermann P. Random Walks, Heat Equation and Distributed Algorithms. *Journal of Computational and Applied Mathematics*, 1994, no. 53, pp. 243–274.
9. Maier R. S. Colliding Stacks: A Large Deviations Analysis. *Random Structures and Algorithms*, 1991, no. 2, pp. 379–421.
10. Sokolov A. V. About Memory Distribution for Two Stacks. *Avtomatizatsiya eksperimenta i obrabotki dannykh*, Petrozavodsk, Karel'skii filial AN SSSR, 1980, pp. 65–71 (In Russian).
11. Aksenova E. A., Sokolov A. V. The Optimal Implementation of Two FIFO-Queues in Single-Level Memory. *Applied Mathematics*, 2011, vol. 2, pp. 1297–1302.
12. Sokolov A. V., Drac A. V. The Linked List Representation of n LIFO-Stacks and/or FIFO-Queues in the Single-Level Memory. *Information Processing Letters*, 2013, vol. 13, pp. 832–835.
13. Aksenova E. A., Drac A. V., Sokolov A. V. Optimal Control of the n FIFO-queues for Infinity Time. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2009, no. 6, pp. 46–54 (In Russian).
14. Drac A. V., Sokolov A. V. Mathematical Analysis of Processing M FIFO-queues. *Stokhasticheskaya optimizatsiya v informatike*, 2012, vol. 8, no. 2, pp. 75–82 (In Russian).
15. Kemeny J. G., Snell J. L. *Finite Markov Chains*. Springer, 1983. 226 p.

МОДЕЛЬ ДЛЯ ОЦЕНКИ СРЕДНЕГО КОЛИЧЕСТВА АКТИВНЫХ АБОНЕНТОВ В СИСТЕМЕ МЕЖМАШИННОГО ВЗАИМОДЕЙСТВИЯ В СЕТЯХ 5-ГО ПОКОЛЕНИЯ

М. А. Гранкин^{а, 1}, аспирант

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Введение: важнейшей характеристикой систем межмашинного взаимодействия (*Machine-to-Machine*) является среднее количество активных абонентов в системе. Под «активным» понимают такого абонента, у которого в данный момент есть готовый для передачи пакет. При большом количестве активных абонентов в сети в системе будет наблюдаться большая средняя задержка на передачу пакета. В 5-м поколении беспроводных сетей предлагаются решения, в которых использование неортогональных ресурсов позволяет реализовать обслуживание большого количества абонентских устройств и обеспечить относительно небольшое количество активных абонентов в сети. При этом вероятность успешной передачи в некотором ресурсе зависит от общего количества передаваемых пакетов в других ресурсах. Данная особенность усложняет необходимый анализ таких систем. **Цель:** разработка модели для анализа характеристик систем межмашинного взаимодействия, построенных на базе сетей 5-го поколения. **Результаты:** предложена новая модель случайного множественного доступа для анализа характеристик систем *Machine-to-Machine* на основе беспроводных сетей последнего поколения. Модель построена с учетом того, что вероятность успешного детектирования пакета от абонентского устройства зависит от общего количества передающих абонентов в системе, даже если они передают в ортогональных каналах. Для анализа таких систем использован метод жидкостной аппроксимации, позволяющий при низких вычислительных затратах оценить такие характеристики, как задержки на передачу пакета и среднее число абонентов, имеющих готовый для передачи пакет. Точность предложенного метода продемонстрирована на численном примере. **Практическая значимость:** предложенный метод и результаты исследования могут быть использованы разработчиками систем межмашинного взаимодействия для оценки среднего энергопотребления и задержек на передачу пакета, в том числе для оперативного перераспределения числа ресурсов, выделяемых для взаимодействия *Machine-to-Machine* систем и систем, ориентированных на передачу данных между людьми (*Human-to-Human*) в гибридных сотовых сетях.

Ключевые слова — межмашинное взаимодействие, жидкостная аппроксимация, случайный множественный доступ, марковский процесс.

Введение

Системы межмашинного взаимодействия (*Machine-to-Machine* — M2M) — одна из ключевых областей в беспроводной связи, используемая в большом диапазоне задач, таких как системы мониторинга, системы «умный» дом (*smart home*), интеллектуальные счетчики (*smart metering*), электронное здравоохранение (*e-health*) и др. Согласно работе [1], сети M2M состоят из датчиков, использующихся для фиксации событий или измерений данных, которые передаются через сети в программное обеспечение для анализа информации. При этом технология M2M удобна для развертывания сенсорных сетей с логической структурой «все-к-одному» и большим количеством датчиков [2].

В отличие от систем, ориентированных на передачу данных между людьми (*Human-to-Human* — H2H), ключевой характеристикой систем M2M

является не скорость передачи данных, а поддержка большого количества абонентов, минимальные задержки на передачу пакета и энергопотребление, которое характеризуется количеством повторных передач и активных абонентов в системе.

В системах M2M большое количество абонентов, имеющих в сети, передают пакеты в пределах небольшого промежутка времени в одном из общих каналов связи [3]. При данном сценарии весьма вероятны ситуации, в которых пакеты абонентов интерферируют друг с другом, что приводит к тому, что ни один из сигналов не может быть успешно принят. Поэтому анализ алгоритмов множественного доступа по такому параметру, как задержка на передачу пакета, является ключевым в сетях, на основе которых могут быть построены системы M2M. С задержками неразрывно связана другая характеристика систем массового обслуживания — среднее количество активных абонентов в системе (данную связь можно иллюстрировать, например, через формулу Литтла). Под «активным» понимают такого абонента, у которого в данный момент есть пакет для передачи. Именно анализу данной характеристики и посвящена большая часть данной статьи.

¹ Научный руководитель — профессор, проректор по науке, доктор технических наук, директор Института информационных систем и защиты информации, заведующий кафедрой безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения Е. А. Крук.

Модели систем случайного множественного доступа

Рассмотрим централизованную беспроводную систему, включающую в себя базовую станцию и большое количество идентичных устройств, которые подключены к базовой станции через общий канал со случайным множественным доступом (СМД).

В настоящее время используются несколько моделей СМД (рис. 1) [4]. В 1970 г. Абрамсон предложил одноканальную модель СМД (рис. 1, а) [5, 6]. В ней представлены ситуации, когда в канале находится несколько одновременно передающих абонентов (обозначим их количество как K).

Модель можно описать, используя следующие допущения.

Допущение 1. Все время разбито на кадры, равные длительности передачи одного пакета. Моменты времени, соответствующие началу кадра, известны всем абонентам. Абонент может начать передачу пакета только в начале очередного кадра. Каждый из абонентов системы, наблюдая выход канала, знает о результате передачи своего пакета.

Допущение 2. В зависимости от значения K будем различать следующие события в канале:

- в канале никто не передает, $K = 0$ (данное событие условно будем называть «пусто»);
- в канале передает только один абонент, $K = 1$ («успех»);
- в канале передают два и более абонентов, $K \geq 2$ («конфликт»).

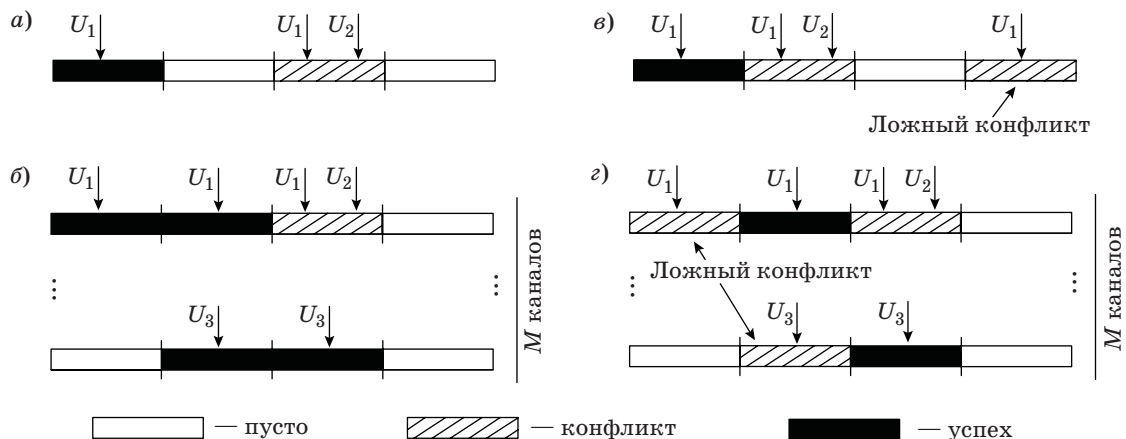
Современные стандарты беспроводных сетей позволяют использовать несколько ортогональных и независимых друг от друга каналов. Исходя из этого была предложена соответствующая модель, которая является обобщением модели Абрамсона [7, 8]. Перед передачей пакета абоненты случайным образом выбирают один из нескольких доступных каналов (рис. 1, б).

«Конфликты» в каналах происходят независимо от событий в других каналах.

Еще одним обобщением одноканальной системы является модель с «ложными конфликтами», предложенная в 1982 г. [9]. В данной модели, даже если пакет передавал один абонент, существовала вероятность того, что он не будет доставлен (рис. 1, в). В работе [9] впервые было введено понятие «ложного конфликта», которое заключается в том, что событие «успех» приемник воспринимает как «конфликт». Этому приводится следующее обоснование: использование помехоустойчивого кодирования дает возможность уменьшить вероятности ошибок вида «пусто»-«успех», «конфликт»-«пусто», «конфликт»-«успех» до пренебрежимо малых значений. Вместе с тем бороться с ошибками вида «успех»-«конфликт» значительно труднее, поскольку множество различных сигналов на входе приемников абонентов при отсутствии шума в ситуации «конфликт» существенно шире, чем в ситуациях «пусто» и «успех». Вследствие этого ошибки типа «ложный конфликт» следует признать в целом ряде случаев более значительными, чем ошибки остальных типов.

В дальнейшем для моделирования сотовых сетей была использована объединенная модель — многоканальная СМД с ложными конфликтами (рис. 1, г) [10]. Таким образом, можно представить следующую схему развития моделей множественного доступа (рис. 2).

Во всех рассмотренных многоканальных моделях вводится допущение о том, что каналы ортогональны, и вероятность успешной передачи не зависит от событий в других каналах. В настоящее время для 5-го поколения беспроводных сетей предлагаются решения, в которых использование неортогональных ресурсов позволяет реализовать обслуживание большого количества абонентских устройств, характерное для систем М2М, и обеспечить относительно небольшое ко-



■ **Рис. 1.** Основные модели СМД: а — одноканальная модель; б — многоканальная модель; в — модель с «ложными конфликтами»; г — многоканальная модель с «ложными конфликтами»

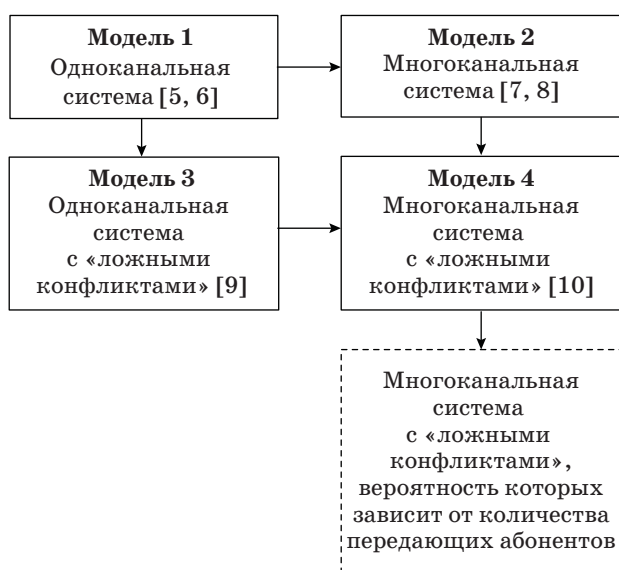


Рис. 2. Взаимосвязь моделей СМД

личество активных абонентов в сети. При этом вероятность успешной передачи в некотором ресурсе зависит от общего количества передаваемых пакетов в других ресурсах. Данная особенность подталкивает к созданию новой модели СМД.

Модель системы СМД с ложными конфликтами, вероятность которых зависит от количества абонентов, передающих в окне

В данной работе предложена новая модель СМД, наиболее подходящая для моделирования сетей М2М, развернутых на базе беспроводных сетей последнего поколения: многоканальная модель с ложными конфликтами, вероятность которых зависит от количества абонентов, ведущих передачу во всех каналах (рис. 3). Данная модель основана на следующей особенности стандартов беспроводных сетей последнего поколения: для передачи пакета абонент сети выбирает одну из доступных преамбул. Если абонент является

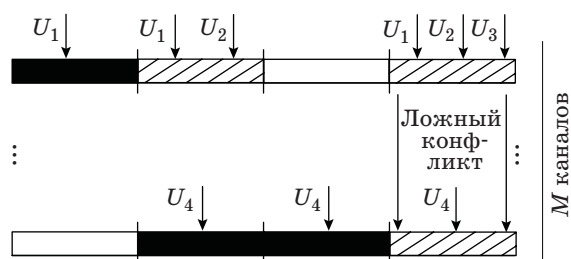


Рис. 3. Предложенная модель СМД с «ложными конфликтами», вероятность которых зависит от количества абонентов, передающих в окне

единственным, кто передает определенную преамбулу, т. е. в канале связи нет конфликтующих с ней сигналов, его преамбула может быть успешно принята. Однако существует вероятность того, что приемник не сможет детектировать данную преамбулу, при этом вероятность детектирования уменьшается с ростом количества одновременно переданных с ней преамбул.

Работу этой системы можно описать как работу многоканальной системы с «ложными конфликтами», вероятность которых $q(i)$ зависит от общего количества абонентов, передающих в окне i . При этом число каналов равняется числу возможных для передачи преамбул.

Основываясь на вышеказанном, можно ввести следующие допущения.

Допущение 3. Помимо трех основных событий, описанных в допущении 1, будем различать событие «ложный конфликт» — такое событие, при котором событие «успех» приемник воспринимает как «конфликт». Вероятность «ложного конфликта» представлена функцией $q(i)$, где i — общее количество абонентов, передающих в данном окне во всех каналах.

Следующий список допущений учитывает рекомендации из работы [3].

Допущение 4. В системе имеются N абонентов, каждый из которых может находиться в одном из двух состояний — «активное» или «неактивное». В начале кадра с вероятностью y абонент переходит из неактивного состояния в активное.

В современных беспроводных сотовых сетях для разрешения конфликтов множественного доступа используются различные вариации алгоритмов экспоненциальной отсрочки или алгоритма «адаптивная Алоха». Для упрощения аналитического анализа будем рассматривать следующий вариант алгоритма Алоха: абонент при появлении пакета передает его сразу в следующем окне, а в случае «конфликта» повторно передает его с вероятностью p . Данный алгоритм, известный как Алоха с немедленной первой передачей (*Immediate First Transmission* — IFT) [11], наиболее близок к основным особенностям механизмов множественного доступа сотовых сетей последнего поколения.

Допущение 5. При появлении пакета абонент передает его, используя один из M независимых каналов. Если в канале, в котором передавал абонент, произошел «успех» и пакет передан, абонент переходит из активного состояния в неактивное. В случае «конфликта» активный абонент повторно передает свой пакет с вероятностью p .

В дальнейшем будем опираться на модель, основанную на этих допущениях, со следующим набором параметров:

N — общее количество абонентов в сети;

y — вероятность появления пакета у абонента в окне;

p — вероятность повторной передачи;
 M — количество каналов (преамбул);
 $q(i)$ — вероятность успешного детектирования,
 где i — количество передающих в окне абонентов.

Оценка среднего количества числа активных абонентов в системе

Введем в рассмотрение случайный процесс $N_{\text{акт}}(t)$ — количество активных абонентов, т. е. абонентов, имеющих готовый для передачи пакет в окне с номером t .

При этом будем рассматривать бесконечно долгое функционирование системы:

$$\bar{N}_{\text{акт}} = \lim_{t \rightarrow \infty} \left(M \left[N_{\text{акт}}^{(t)} \right] \right). \quad (1)$$

Величину $\bar{N}_{\text{акт}}$ будем далее называть средним числом активных абонентов. Непосредственно из введенных выше допущений модели следует, что $N_{\text{акт}}(t)$ является односвязной марковской цепью с $N + 1$ состояниями, включая состояние, когда в системе нет активных абонентов. В данной марковской цепи всегда существует стационарное (финальное) распределение вероятностей состояний $\{P_i\}$, где P_i — вероятности пребывания в i -м состоянии, что позволяет вычислить среднее количество активных абонентов в системе следующим образом:

$$\bar{N}_{\text{акт}} = \sum_{i=1}^N i P_i.$$

Стационарное распределение может быть найдено решением системы из $N + 1$ линейных уравнений. Частный случай для такой цепи для модели, в которой отсутствуют «ложные конфликты», приведен в работе [11].

Как отмечено выше, одной из особенностей сетей межмашинного взаимодействия является большое количество абонентов в сети. Данная особенность делает поиск стационарного распределения вычислительно сложным. Например, метод решения системы линейных уравнений Гаусса требует $O(N^3)$ арифметических операций. Если задать количество абонентов как $N = 10^6$, компьютер, осуществляющий миллиард операций в секунду, будет решать данную систему для одного набора параметров больше десяти дней.

Покажем, как $\bar{N}_{\text{акт}}$ можно оценить со значительно меньшими вычислительными затратами. В работе [12] предложена вычислительно простая приближенная оценка одноканальной системы Алоха — метод жидкостной аппроксимации. Далее мы обобщим данную оценку для случая многоканальной системы с «ложными конфликтами», вероятность которых зависит от общего количества передаваемых пакетов. Излагаемый авторами [12] метод основан на замене случайно-

го процесса $N_{\text{акт}}(t)$, представляющего собой количество активных абонентов в момент времени t , марковским процессом диффузионного типа. Для описания данного метода введем в рассмотрение следующие величины:

$X^{(t)}$ — число пакетов, появляющихся в системе в некотором окне t ;

$Y^{(t)}$ — число пакетов, успешно переданных в окне t .

Рассмотрим условные математические ожидания данных величин:

$M \left[X^{(t)} \mid N_{\text{акт}}^{(t)} = n \right]$ будем называть интенсивностью входного потока (обозначим как $\lambda_{\text{вх}}(n)$);

$M \left[Y^{(t)} \mid N_{\text{акт}}^{(t)} = n \right]$ будем называть интенсивностью выходного потока (обозначим как $\lambda_{\text{вых}}(n)$).

В работе [12] показано, что если уравнение

$$\lambda_{\text{вх}}(n) = \lambda_{\text{вых}}(n) \quad (2)$$

имеет единственный корень (обозначим его как n_0), то справедливо следующее приближенное равенство:

$$\bar{N}_{\text{акт}} \approx n_0.$$

Погрешность вычисления данного метода не превышает 2 % для случая одноканальной системы без «ложных конфликтов» при том, что из-за простоты реализации она позволяет проанализировать сценарии с большим количеством абонентов, характерным для систем М2М [12].

Однако недостатком данного метода можно считать нереализуемость его при некоторых параметрах системы — когда уравнение (2) имеет несколько корней или не имеет корней вообще. Покажем, что данный недостаток несущественен при анализе предложенной модели системы М2М. Для того чтобы это показать, рассмотрим понятие «стабильной» системы.

Интенсивность входного потока $\lambda_{\text{вх}}(n)$ убывает с увеличением n линейно до нуля, что соответствует ситуации, когда все абоненты в системе являются активными. В зависимости от того, сколько пересечений имеют функции $\lambda_{\text{вых}}(n)$ и $\lambda_{\text{вх}}(n)$, что в свою очередь зависит от параметров системы, можно различать три возможных состояния системы (рис. 4) [13]:

— «стабильная» — если имеется одно пересечение между $\lambda_{\text{вых}}(n)$ и $\lambda_{\text{вх}}(n)$;

— «бистабильная» — если имеется несколько пересечений между $\lambda_{\text{вых}}(n)$ и $\lambda_{\text{вх}}(n)$;

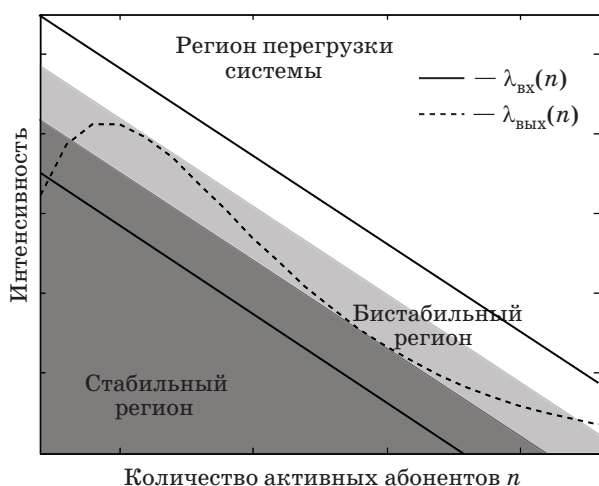
— «перегруженная» — если $\lambda_{\text{вых}}(n) < \lambda_{\text{вх}}(n)$ для любых n .

Для практики представляют интерес только стабильные системы, в которых наиболее вероятное значение $N_{\text{акт}}$ близко к $\bar{N}_{\text{акт}}$, существенно меньше N и почти всегда является единственным максимумом распределения вероятности.

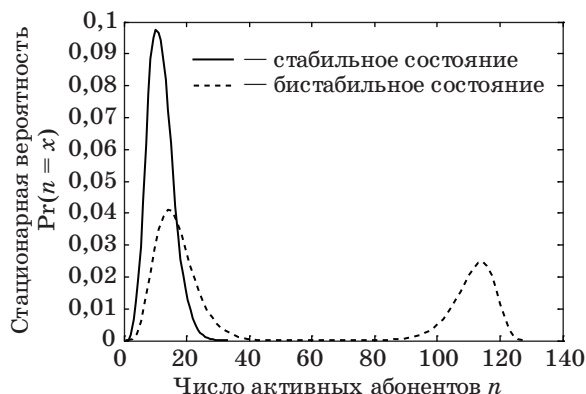
В случае бистабильной системы таких локальных максимумов два, при этом один из них соответствует низкой загрузке системы и высокой вероятности успешной передачи (или, иначе говоря, большой интенсивности выходного потока), а второй — неприемлемо высокой загрузке и практически нулевой вероятности успеха. Также при некоторых наборах параметров может возникнуть ситуация «перегрузки», в которой $N_{\text{акт}} \rightarrow N$.

Для наглядности можно построить зависимость стационарной вероятности пребывания в стабильном и бистабильном состоянии марковского процесса $N_{\text{акт}}$ (рис. 5).

Для использования на практике бистабильные и перегруженные состояния не представляют интереса для М2М-систем, так как могут привести к большому количеству активных абонентов. Поэтому в рамках этой статьи будут рассмотрены только стабильные системы, тем самым нивели-



■ Рис. 4. Функции интенсивностей для стабильных и нестабильных систем



■ Рис. 5. Распределение вероятностей пребывания по состояниям марковской цепи для системы в стабильном и бистабильном состоянии

руя недостатки метода жидкостной аппроксимации для анализа данной модели.

Непосредственно из сформулированного выше определения входного потока и допущения 3 справедливо следующее утверждение.

Утверждение 1. Интенсивность входного потока для модели СМД системы межмашинного взаимодействия, если в системе n активных абонентов, равна

$$\lambda_{\text{вх}}(n) = y(N - n).$$

Перейдем к рассмотрению интенсивности выходного потока. Частный случай интенсивности выходного потока для многоканальной системы без «ложных конфликтов» рассмотрен в работе [14]. Для вывода формулы к предложенной модели можно привести следующие рассуждения.

Если в системе имеется M каналов, то количество успешно переданных пакетов в системе в момент времени $Y^{(t)}$ будет равно сумме количества успешно переданных пакетов $Y_i^{(t)}$ в каждом из M каналов:

$$Y^{(t)} = \sum_{i=1}^M Y_i^{(t)}.$$

Вычислим математическое ожидание левой и правой частей равенства:

$$M[Y^{(t)} | N_{\text{акт}}^{(t)} = n] = \sum_{i=1}^M M[Y_i^{(t)} | N_{\text{акт}}^{(t)} = n].$$

Среднее количество успешно переданных пакетов в канале равно вероятности события «успех» и не зависит от того, в каком конкретно канале передает абонент. Таким образом:

$$\lambda_{\text{вых}}(n) = M \cdot \text{Pr}\{\text{усп. в одном кан.} | n\}.$$

Для начала рассмотрим систему, в которой вероятность «ложного конфликта» равняется нулю, соответственно вероятность детектирования $q(i) = 1$ для любых $i \in 1 : N$. Сложное событие, связанное с успешной передачей пакета в слоте, является объединением двух простых событий. Первое из них состоит в том, что ни один из n абонентов, имеющих готовый для передачи пакет, не передает, но при этом возникает ровно один новый пакет у $N - n$ неактивных абонентов, который передается в этом окне. Второе событие соответствует передаче пакета ровно одним из активных абонентов и отсутствию новых пакетов у остальных абонентов:

$$\begin{aligned} \text{Pr}\{\text{усп. в одном кан.} | n\} &= \\ &= \text{Pr}\{\text{нет новых аб.}\} \times \\ &\times \text{Pr}\{\text{передал один активный аб.}\} + \\ &+ \text{Pr}\{\text{появился новый аб.}\} \times \\ &\times \text{Pr}\{\text{активные аб. не передали}\}. \end{aligned}$$

Рассмотрим конкретный из M каналов системы: вероятность того, что какой-нибудь из активных абонентов будет передавать в конкретном из M каналов, равна произведению вероятности того, что будет передавать активный абонент, — p и вероятности того, что абонент выберет именно этот канал, — $1/M$. Так как эти события независимы, вероятность того, что активный абонент будет передавать в данном канале, равняется $\frac{p}{M}$, и, очевидно, вероятность того, что активный абонент в данном канале передавать не будет: $1 - \frac{p}{M}$.

Те же рассуждения справедливы и для вероятности того, что в данном канале будет передавать неактивный абонент. Следовательно, вероятность успеха в одном из M каналов

$$\begin{aligned} \Pr\{\text{усп. в одном кан.} | n\} &= \\ &= n \frac{p}{M} \left(1 - \frac{p}{M}\right)^{n-1} \left(1 - \frac{y}{M}\right)^{N-n} + \\ &+ \left(1 - \frac{p}{M}\right)^n (N-n) \frac{y}{M} \left(1 - \frac{y}{M}\right)^{N-n-1}. \end{aligned}$$

Для тривиального случая, когда вероятность «ложного конфликта» $q(i) = c$, т. е. не зависит от того, сколько абонентов передают в окне, итоговая формула выглядит следующим образом:

$$\begin{aligned} \Pr\{\text{усп. в одном кан.} | n, q(n)\} &= \\ &= c \cdot \Pr\{\text{усп. в одном кан.} | n\}. \end{aligned}$$

Вернемся к рассмотрению случая, когда вероятность «ложного конфликта» зависит от количества передающих в окне абонентов. Сложное событие, связанное с успешной передачей пакета в слоте, является объединением N простых событий. Каждое из них состоит в том, что в канале абонент передавал один и пакет был успешно детектирован с вероятностью $q(i)$, где i — общее количество передающих в окне абонентов:

$$\begin{aligned} \Pr\{\text{усп. в одном кан.} | n\} &= \\ &= \sum_{i=1}^N q(i) \cdot \Pr\{i \text{ аб. передают} | n\} \times \\ &\times \Pr\{1 \text{ аб. в канале}\}. \end{aligned} \quad (3)$$

Вероятность того, что в определенном канале передает один абонент, — это вероятность того, что один из i абонентов выбрал данный канал, а остальные $i-1$ абонентов передают в других $M-1$ каналах:

$$\Pr\{1 \text{ аб. в канале}\} = i \left(\frac{1}{M} \right) \left(1 - \frac{1}{M} \right)^{i-1}.$$

Рассмотрим вероятность $\Pr\{i \text{ аб. передают}\}$ того, что i абонентов передают в данном окне.

Введем случайную величину $K(n)$ — количество передающих абонентов в окне, при условии, что в системе имеется n активных абонентов. Для этой случайной величины непосредственно из допущений 4 и 5 вытекает следующее утверждение.

Утверждение 2:

$$K(n) = K_a(n) + K_n(n),$$

где $K_a(n)$ — количество абонентов, перешедших из неактивного состояния в активное (согласно допущению 4, данные абоненты будут передавать в этом окне); $K_n(n)$ — количество активных абонентов, принявших решение осуществить повторную передачу. Случайные величины $K_a(n)$ и $K_n(n)$ независимы и распределены по биномиальному закону.

Из утверждения 2 следует, что распределение вероятностей случайной величины $K(n)$ равно дискретной свертке распределения вероятностей случайных величин $K_a(n)$ и $K_n(n)$:

$$\begin{aligned} \Pr\{i \text{ передают} | n\} &= \\ &= \sum_{j=\max(0, i-(N-n))}^{\min(n, i)} C_n^j p^j (1-p)^{n-j} \times \\ &\times C_{N-n}^{i-j} y^{i-j} (1-y)^{N-n-i+j}. \end{aligned}$$

При большом количестве абонентов N вычисление данной суммы напрямую ведет к неприемлемым вычислительным затратам. Покажем, каким образом можно приближенно вычислить данную вероятность.

Как уже говорилось, системы М2М характерны большим количеством абонентов в системе N . Это приводит к тому, что в зависимости от количества активных абонентов числа n , или $N-n$, или оба сразу будут достаточно велики. В системах М2М низкая интенсивность входного потока, т. е. вероятность y , с которой абонент переходит из неактивного состояния в активное. При этом, чтобы снизить вероятность «конфликта» в системе с большим количеством абонентов, необходимо снизить вероятность передачи p .

В этих условиях распределения случайных величин $K_a(n)$ и $K_n(n)$ можно аппроксимировать случайными величинами $I_a(n)$ и $I_n(n)$, распределенными по пуассоновскому закону с параметрами $\lambda_a(n) = pn$ и $\lambda_n(n) = y(N-n)$ соответственно. Согласно неравенству Ле Кама [15], ошибка аппроксимации в этом случае не будет превышать величины $4p$ (или $4y$ соответственно), т. е.

$$\begin{aligned} \sum_{k=0}^{\infty} |\Pr\{K_a(n) = k\} - \Pr\{I_a(n) = k\}| &< 4p; \\ \sum_{k=0}^{\infty} |\Pr\{K_n(n) = k\} - \Pr\{I_n(n) = k\}| &< 4y. \end{aligned}$$

Сумма независимых пуассоновских случайных величин также имеет распределение Пуассона, следовательно:

$$\Pr\{i \text{ передают} | n\} \approx \frac{\lambda_{\text{сум}}^i(n)}{i!} e^{-\lambda_{\text{сум}}(n)},$$

где $\lambda_{\text{сум}}(n) = \lambda_{\text{н}}(n) + \lambda_{\text{а}}(n) = y(N - n) + pn$.

Таким образом, вероятность «успеха» в одном канале при условии, что в системе имеются n активных абонентов, может быть аппроксимирована следующим выражением:

$$\begin{aligned} \Pr\{\text{усп. в одном кан.} | n\} &\approx \left(\sum_{i=1}^N q(i) \frac{\lambda_{\text{сум}}^i}{i!} e^{-\lambda_{\text{сум}}} i \frac{1}{M} \left(1 - \frac{1}{M}\right)^{i-1} \right) = \\ &= \frac{1}{M} \left(1 - \frac{1}{M}\right)^{-1} \sum_{i=1}^N q(i) i \frac{\left(\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)\right)^i}{i!} e^{-\lambda_{\text{сум}}} = \\ &= \frac{1}{M} \left(1 - \frac{1}{M}\right)^{-1} e^{-\frac{\lambda_{\text{сум}}}{M}} \times \\ &\times \sum_{i=1}^N q(i) i \frac{\left(\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)\right)^i}{i!} e^{-\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)}. \quad (4) \end{aligned}$$

Рассмотрим сумму более подробно. Так как N большое, данное выражение является аппроксимацией математическим ожиданием от произведения случайных величин, одно из которых распределено по пуассоновскому закону (обозначим его как I) с параметром $\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)$:

$$M[q(i) \cdot I] = \sum_{i=1}^N q(i) i \frac{\left(\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)\right)^i}{i!} e^{-\lambda_{\text{сум}} \left(1 - \frac{1}{M}\right)}.$$

Таким образом, можно сформулировать следующие утверждения.

Утверждение 3. Интенсивность выходного потока для модели СМД системы межмашинного взаимодействия, если в системе n активных абонентов, равна

$$\begin{aligned} \lambda_{\text{вых}}(n) &= \\ &= M \sum_{j=\max(0, i-(N-n))}^{\min(n, i)} \left(C_n^j p^j (1-p)^{n-j} C_{N-n}^{i-j} \times \right. \\ &\quad \left. \times y^{i-j} (1-y)^{N-n-i+j} \right) \end{aligned}$$

и может быть оценена как

$$\begin{aligned} \hat{\lambda}_{\text{вых}}(n) &= \left(1 - \frac{1}{M}\right)^{-1} e^{-\frac{\lambda_{\text{сум}}(n)}{M}} \times \\ &\times \sum_{i=1}^N q(i) i \frac{\left(\left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n)\right)^i}{i!} e^{-\left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n)}, \end{aligned}$$

где $\lambda_{\text{сум}}(n) = y(N - n) + pn$.

Для иллюстрации оценки, полученной данной формулой, рассмотрим случай, когда зависимость $q(i)$ представляется простейшей элементарной функцией. При этом для какой-либо сложной зависимости $q(i)$ требуется подобрать функцию, которая будет нижней оценкой для такой зависимости. Данное указание можно объяснить следующим логическим рассуждением: в системах М2М целесообразно оценивать максимальную нагрузку на сеть для вычисления пикового значения энергопотребления в сети. Из этого следует, что требуется найти верхнюю оценку для количества активных абонентов и, соответственно, нижнюю оценку вероятности события «успех», которая прямо пропорциональна вероятности детектирования.

Рассмотрим простейший случай, когда вероятность «ложного конфликта» не зависит от количества абонентов $q(i) = c$. Тогда

$$M[q(i) \cdot I] = M[cI] = c \left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n),$$

и в этом случае формула (3) сводится к простому выражению

$$\begin{aligned} \Pr\{\text{усп. в одном кан.} | n\} &\approx \\ &\approx \frac{c}{M-1} e^{-\frac{\lambda_{\text{сум}}(n)}{M}} \left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n). \end{aligned}$$

Рассмотрим более сложный случай, когда функция $q(i)$ — линейная функция вида $q(i) = ik + c$. Таким образом, вычисление формулы сводится к следующему выражению:

$$\begin{aligned} M[q(i) \cdot I] &= M[(kI + c) \cdot I] = kM[I^2] + cM[I] = \\ &= k \left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n) (1 + \lambda_{\text{сум}}(n)) + c \left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n) = \\ &= \lambda_{\text{сум}}(n) \left(1 - \frac{1}{M}\right) \left(c + k + k \left(1 - \frac{1}{M}\right) \lambda_{\text{сум}}(n)\right). \quad (5) \end{aligned}$$

Все предыдущие выкладки были выведены с учетом того допущения, что $q(i) \geq 0$ для любых

$i \in (1, \infty)$. Следует дополнить формулу (5), рассматривая более общий случай:

$$\begin{cases} q(i) \geq 0, & \text{если } i \in (1, N_{\max}) \\ q(i) = 0, & \text{если } i > N_{\max} \end{cases},$$

где $N_{\max}$ — максимальное количество передающих абонентов, при котором вероятность успешной передачи не равняется нулю. $N_{\max} = -\frac{c}{k}$ для

функции вида $q(i) = ik + c$. В этом случае для усеченной плотности вероятности необходимо вычислить нормирующий коэффициент, равный $1/F_{Pois}(I, N_{\max})$, где $F_{Pois}(I, x)$ — функция распределения пуассоновской случайной величины с параметром I . На основе всего вышесказанного и формул (4) и (5), наконец, можно сформировать следующее утверждение.

Утверждение 4. Интенсивность выходного потока для модели СМД системы межмашинного взаимодействия, если в системе n активных абонентов и вероятность успешного детектирования — линейно убывающая функция вида $q(i) = ik + c$, может быть оценена как

$$\hat{\lambda}_{\text{вых}}(n) = \frac{\lambda_{\text{сум}}(n)}{F_{Pois}\left(\lambda_{\text{сум}}, -\frac{c}{k}\right)} e^{-\frac{\lambda_{\text{сум}}(n)}{M}} \times \left(c + k + k\left(1 - \frac{1}{M}\right)\lambda_{\text{сум}}(n)\right), \quad (6)$$

где $\lambda_{\text{сум}}(n) = y(N - n) + pn$.

Поскольку рассматриваются параметры системы, в которых система стабильна, или, иначе говоря, параметры, при которых уравнение (2) имеет один корень в интервале $n \in [0, N]$, то для решения данного уравнения численным методом можно воспользоваться методом дихотомии. Задача заключается в том, чтобы найти и уточнить этот корень методом половинного деления. Другими словами, требуется найти приближенное значение корня с заданной точностью до единицы. Легко показать, что в этом случае сложность поиска оценки предела (1) через поиск корня уравнения (2) не превышает сложность $\log_2(N)$, что позволяет ее использовать как «быструю» оценку среднего количества абонентов в системе.

Численный пример

Для демонстрации точности результатов можно взять один из типовых сценариев для моделирования сетей M2M [3]. Требуется определить среднее количество активных абонентов в сети $N_{\text{акт}}$ для следующих параметров:

— вероятность появления пакета у абонента в окне $y = 10^{-3}$;

— вероятность повторной передачи $p = 10^{-2}$;

— количество абонентов $N = 500 \div 5000$;

— количество каналов (преамбул) $M = 64$;

— вероятность успешного детектирования убывает линейно и задается выражением

$$\begin{cases} q(i) = -9 \cdot 10^{-4} + 0,9, & i \in (1, 1000) \\ q(i) = 0, & i > 1000 \end{cases}.$$

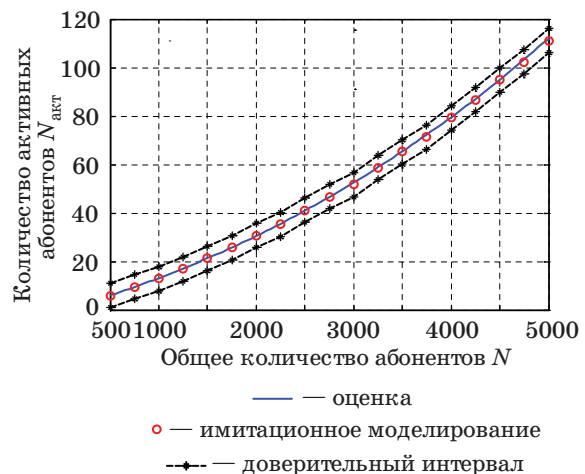
Для численного примера был выбран сценарий, для которого уже вычислительно сложно получить точный результат методом, основанным на вычислении стационарного распределения марковской цепи. Поэтому для демонстрации точности предложенной оценки было использовано имитационное моделирование для заданного доверительного интервала $\varepsilon > |\bar{N}_{\text{акт}} - \hat{N}_{\text{акт}}|$, равного $\varepsilon = 5$ абонентов, где $\hat{N}_{\text{акт}}$ — оценка, полученная методом имитационного моделирования. Количество экспериментов для достоверной оценки моделирования было получено из неравенства Чебышева:

$$num = \left\lceil \frac{t_{\alpha}^2 \hat{\sigma}^2}{\varepsilon^2} \right\rceil,$$

где t_{α}^2 — квантиль нормального распределения, а верхняя оценка дисперсии выборки вычислялась как $\hat{\sigma}^2 = N/4$.

Для имитационного моделирования системы с общим количеством абонентов, равным (1000, 1500, 2000, 3000, 5000), было проведено (37, 102, 201, 498, 1482) · 10³ экспериментов соответственно.

Как видно из результатов эксперимента (рис. 6), погрешность оценки не выходит за рамки доверительного интервала.



■ Рис. 6. Результаты эксперимента

Закключение

В статье предложена новая модель случайного множественного доступа, отражающая особенности систем межмашинного взаимодействия, построенных на базе беспроводных сетей 5-го поколения. Модель учитывает зависимость вероятности успешного детектирования от числа одновременно передающих абонентов. Для оценки среднего числа активных абонентов предложена вычислительно простая

приближенная оценка методом жидкостной аппроксимации. При малой вычислительной сложности данная оценка обладает высокой точностью.

В дальнейших исследованиях планируется рассмотреть возможности использования предложенной модели и методики расчета ее характеристик для решения задачи оперативного перераспределения числа ресурсов, выделяемых для Machine-to-Machine и Human-to-Human взаимодействия в гибридных системах.

Литература

1. Emmerson B. M2M: the Internet of 50 Billion Devices // Win-Win Magazine. 2010. N 1. P. 19–22.
2. Шепета А. П., Евсеев Г. С., Бакин Е. А. Нижняя граница длительности периода сбора информации в сенсорной сети // Информационно-управляющие системы. 2011. № 6(55). С. 64–67.
3. GPP TR 37.868. RAN Improvements for Machine-type Communications. Ver. 1.0.0. Oct. 2011. <http://www.3gpp.org/DynaReport/37868.htm> (дата обращения: 01.06.2013).
4. Chan D., Berger T. Upper Bound for the Capacity of Multiple Access Protocols on Multipacket Reception Channels // IEEE Intern. Symp. on Information Theory Proc., Cambridge, 1–6 July 2012. P. 1603–1607.
5. Abramson N. The Aloha System: Another Alternative for Computer Communications // Fall Joint Computer Conf. Ser. AFIPS'70 (Fall): Proc. 17–19 Nov. 1970. New York: ACM, 1970. P. 281–285.
6. Bertsekas D., Gallager R. Data Networks. — Prentice Hall, 1992. — 556 p.
7. Mehravari N. Random-Access Communication with Multiple Reception // IEEE Trans. Inf. Theory. 1990. Vol. 36. N 3. P. 614–622.
8. Pountourakis I., Sykas E. Analysis, Stability and Optimization of Aloha-type Protocols for Multichannel Networks // Computer Communications. Dec. 1992. Vol. 15. Iss. 10. P. 619–629.
9. Евсеев Г. С., Ермолаев Н. Г. Оценки характеристик разрешения конфликтов в канале со свободным доступом и шумом // Проблемы передачи информации. 1982. № 18. С. 101–105.
10. Евсеев Г. С., Тюрликов А. М. Алгоритмы случайного доступа к системе параллельных каналов с зависимым шумом // X Всес. школа-семинар по вычислительным сетям. Тбилиси, 1985. № 2.
11. Yue W., Yutaka M. Performance Analysis of Multi-Channel and Multi-Traffic on Wireless Communication Networks. — Springer US, 2012. — 324 p.
12. Авен О. И., Гурий Н. Н., Коган Я. А. Оценка качества и оптимизация вычислительных систем. — М.: Наука, 1982. — 464 с.
13. Lam S., Kleinrock L. Packet Switching in a Multi-access Broadcast Channel: Dynamic Control Procedures // IEEE Transactions on Communications. Sept. 1975. Vol. COM-23. P. 891–904.
14. Galinina O., Turlikov A., Andreev S., Koucheryavy Y. Stabilizing Multi-Channel Slotted Aloha For Machine-type Communications // IEEE Intern. Symp. on Information Theory, Istanbul, 7–12 July 2013. P. 2119–2123.
15. Neyman J., Cam L. On the Distribution of Sums of Independent Random Variables // Bernoulli, Bayes, Laplace: Anniversary. New York: Springer-Verlag, 1965. P. 179–202.

UDC 004.728.3.057.4

doi:10.15217/issn1684-8853.2015.5.72

Model for Estimating the Average Number of Backlogged Users of M2M Systems in 5G Mobile Networks

Grankin M. A.^a, Post-Graduate Student, m.a.grankin@gmail.com

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

Purpose: The average number of backlogged users is the most important characteristic of Machine-to-Machine communication. A backlogged user is a user with a packet to transmit. A larger number of backlogged users leads to a bigger packet transmission delay. There are solutions for the 5th generation wireless networks which use non-orthogonal resources. This technique allows servicing of a large number of users' devices, providing a relatively small number of backlogged users in the network. The probability of packet detection in a channel depends on the total number of packets transmitted in other resources. This feature complicates the necessary analysis of such systems. The purpose of this work is developing a model for the analysis of Machine-to-Machine systems based on 5th generation wireless networks. **Results:** A new model of random multiple access has been proposed for the analysis of Machine-to-Machine systems based on 5th generation wireless networks. This model takes into account that the packet detection probability

depends on the total number of transmitting users in the system, even if they transmit in orthogonal channels. To analyze such systems, fluid approximation technique was used. With a low computational cost, this technique allows you to estimate such characteristics of Machine-to-Machine systems as packet transmission delays or the average number of backlogged users. The accuracy of the proposed method is demonstrated by a numerical example. **Practical relevance:** The proposed method and the research results can be used by developers of Machine-to-Machine systems to estimate the average power consumption and transmission delays, including real-time reallocation of the available resources in Machine-to-Machine and Human-to-Human hybrid systems.

Keywords — Machine-to-Machine, Machine-Type Communications, Fluid Approximation, Random Multiple Access, Markov Process.

References

1. Emmerson B. M2M: the Internet of 50 Billion Devices. *Win-Win Magazine*, 2010, no. 1, pp. 19–22.
2. Shepeta A. P., Evseev G. S., Bakin E. A. Low Bound of Frame Duration in Convergecast Sensor Network. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2011, no. 6(55), p. 64–67 (In Russian).
3. GPP TR 37.868. RAN Improvements for Machine-type Communications. Ver. 1.0.0. October 2011. Available at: <http://www.3gpp.org/DynaReport/37868.htm> (accessed 1 June 2013).
4. Chan D., Berger T. Upper Bound for the Capacity of Multiple Access Protocols on Multipacket Reception Channels. *IEEE Intern. Symp. on Information Theory Proc.*, Cambridge, 1–6 July 2012, pp. 1603–1607.
5. Abramson N. The Aloha System: Another Alternative for Computer Communications. *Proc. of the November 17–19, 1970, Fall Joint Computer Conf., ser. AFIPS'70 (Fall)*. New York: ACM, 1970, pp. 281–285.
6. Bertsekas Dimitri, Gallager Robert. *Data Networks*. Prentice Hall, 1992. 556 p.
7. Mehravari N. Random-Access Communication with Multiple Reception. *IEEE Transactions on Information Theory*, 1990, vol. 36, no. 3, pp. 614–622.
8. Pountourakis I., Sykas E. Analysis, Stability and Optimization of Aloha-type Protocols for Multichannel Networks. *Computer Communications*, Dec. 1992, vol. 15, iss. 10, pp. 619–629.
9. Evseev G. S., Ermolaev N. G. Performance Evaluation of the Collision Resolution for a Random-Access Noisy Channel. *Problemy peredachi informatsii* [Problems Information Transmission], 1982, vol. 18, iss. 2, pp. 101–105 (In Russian).
10. Evseev G. S., Turlikov A. M. Algorithms for Random Access to a System of Parallel Channels with Dependent Noise. *X Vsesoiuznaia shkola-seminar po vychislitel'nykh setiam* [X All-Union Workshop on Data Network], Tbilisi, 1985, no. 2 (In Russian).
11. Yue W., Yutaka M. *Performance Analysis of Multi-Channel and Multi-Traffic on Wireless Communication Networks*. Springer US, 2012. 324 p.
12. Aven, O. I., Gurin N. N., Kogan Y. A. *Otsenka kachestva i optimizatsiia vychislitel'nykh sistem* [Performance Evaluation and Optimization of Computer Systems]. Moscow, Nauka Publ., 1982. 464 p. (In Russian).
13. Lam S., Kleinrock L. Packet Switching in a Multicast Broadcast Channel: Dynamic Control Procedures. *IEEE Transactions on Communications*, September 1975, vol. COM-23, pp. 891–904.
14. Galinina O., Turlikov A., Andreev S., Koucheryavy Y. Stabilizing Multi-Channel Slotted Aloha for Machine-type Communications. *IEEE International Symposium on Information Theory*, Istanbul, 7–12 July 2013, pp. 2119–2123.
15. Neyman J., Cam L. On the Distribution of Sums of Independent. *Bernoulli, Bayes, Laplace: Anniversary*, New York, Springer-Verlag, 1965, pp. 179–202.

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы зарегистрируетесь на сайте НЭБ (<http://elibrary.ru/defaultx.asp>), будет создана Ваша личная страничка, содержание которой составят не только Ваши персональные данные, но и перечень всех Ваших печатных трудов, имеющих в базе данных НЭБ, включая диссертации, патенты и тезисы к конференциям, а также сравнительные индексы цитирования: РИНЦ (Российский индекс научного цитирования), h (индекс Хирша) от Web of Science и h от Scopus. После создания базового варианта Вашей персональной страницы Вы получите код доступа, который позволит Вам редактировать информацию, помогая создавать максимально объективную картину Вашей научной активности и цитирования Ваших трудов.

РАЗРАБОТКА ЯЗЫКА УПРАВЛЕНИЯ КОСМИЧЕСКИМИ АППАРАТАМИ

Космынина Н. А.^а, инженер

Легалов А. И.^б, доктор техн. наук, профессор

Барков А. В.^а, канд. техн. наук, начальник сектора

Лапин А. А.^а, канд. техн. наук, начальник сектора

^аАО «Информационные спутниковые системы» им. академика М. Ф. Решетнёва», Железногорск, Красноярский край, РФ

^бСибирский федеральный университет, Красноярск, РФ

Постановка проблемы: в настоящее время возможности языка управления космическими аппаратами, эксплуатируемого на базе акционерного общества «Информационные спутниковые системы» им. академика М. Ф. Решетнёва, перестали соответствовать требованиям предметной области: язык не поддерживает новых возникающих функциональных задач и слишком сложен для быстрого освоения, что вызвало необходимость разработки нового языка управления. **Результаты:** на основании анализа ограничений применяемого языка управления, требований предметной области и существующих аналогов был спроектирован и разработан проблемно-ориентированный язык управления «Дельта», обладающий простым синтаксисом, с применением терминов предметной области в качестве ключевых слов, поддерживающий как русский, так и английский варианты написания операторов языка. В языке поддерживаются основные функции управления: выдача управляющих воздействий, анализ поступающей телеметрии, паузы, печать сообщений, условные операторы и др. **Практическая значимость:** данная разработка предназначена для управления большинством космических аппаратов производства акционерного общества «Информационные спутниковые системы» как в отечественных, так и в международных проектах.

Ключевые слова — предметно-ориентированный язык, проблемно-ориентированный язык, управление космическими аппаратами, процедура управления.

Введение

В настоящее время возможности, предоставляемые использованием космических систем, применяются в различных областях: связи, навигации, геодезии, телевидении и др. Для поддержания выполнения системой указанных целевых функций необходимо обеспечить решение задач управления космическими аппаратами (КА): удержание на заданной орбите, включение/отключение различных режимов работы, изменение алгоритма функционирования полезной нагрузки и др., — что решается выдачей соответствующих команд с Земли на КА [1, 2]. Логическая совокупность команд, выдаваемая на борт КА для реализации единичной задачи управления, называется процедурой управления. Проблемно-ориентированные языки позволяют описать процедуру управления в виде сценария и выполнять его в ходе сеансов связи с КА в автоматизированном режиме неограниченное количество раз. Примеры таких языков: Python, PLUTO, CSL, STOL, TAO, CSTOL, CCL, JAS, ICL, Btscript, Cecil, CIL, TOPE/tcl, UCL, Elisa, PIL, ЯОТР и др.

Обзор и сопоставление ряда языков управления КА

Функциональные возможности языков управления КА обоснованы необходимостью решения

задач управления КА и включают, но не ограничиваются, следующим набором:

- выдача команд управления на КА;
- анализ текущих значений телеметрических (ТМ) параметров на соответствие допустимому интервалу значений;
- сообщение оператору;
- паузы между командами;
- вызов другой процедуры управления из данной;
- комментарии к тексту процедуры.

В связи с большим разнообразием существующих языков управления [3] производятся попытки унификации, для чего разработаны следующие документы: стандарт ECSS-E-ST-70-32C: Test and Operations Procedure Language (PLUTO) [4]; спецификация OMG SOLM [5], — содержащие описание возможностей, которыми, по мнению разработчиков, должен удовлетворять любой язык управления для обеспечения совместности различных систем управления.

Рассмотрим некоторые применяемые языки.

Язык Python [6] — переносимый, высокоуровневый, объектно-ориентированный универсальный язык, поддерживающий параллельное выполнение сценариев управления КА. Применяется в специальном программном обеспечении управления (СПО-У) GSEOS, разработанном компанией I&T, США. Помимо этого, язык Python применяется для автоматизации управления КА

с помощью СПО-У «hifly» [7], разработанном компанией GMV, Испания.

Язык PLUTO [8], применяющийся в СПО-У SCOS Mission Control System, разработанном компанией SciSys (Англия, Германия), а также в СПО-У STEPS, разработанном институтом INPE, Бразилия, — высокоуровневый язык программирования, поддерживающий циклы, арифметические операции, планирование выполнения. Процедура управления, созданная с использованием данного языка, должна содержать следующие разделы: раздел объявлений; раздел условий старта выполнения процедуры (по наступлению конкретного времени или выполнению ряда условий); тело процедуры; раздел проверок оборудования (если в ходе выполнения данного раздела произошел сбой, то выполнение процедуры останавливается); раздел постусловий (проверка того, что процедура была выполнена правильно).

Язык SCL применяется для управления КА в СПО-У ASPEN, компания Jet Propulsion Laboratory [9], США, а также в СПО-У SCL [10], компания Interface & Control Systems, Inc, США. На SCL можно создавать как скрипты управления, так и правила управления. Правила предназначены для мониторинга внештатных ситуаций в подсистемах КА, при появлении которых предусматривается автоматический запуск соответствующего скрипта. Скрипты управления предназначены для исполнения в ходе сеанса связи с КА для выполнения какой-либо целевой задачи. Скрипты могут выполняться параллельно, на основе приоритетов, назначаемых при создании скрипта. Кроме функций, общих для всех языков управления, SCL поддерживает циклы и выражения присваивания.

Компанией NASDA, Япония [11], разработан язык управления КА, предназначенный для формирования и исполнения процедур управления в составе СПО-У SMACS. Помимо общих функций языка управления, данный язык поддерживает обработку следующих событий: изменение конфигурации сети, смену используемой для управления КА станции. Процедура также может производить запись телеметрических данных и сигнализировать о возникших аномалиях.

Компанией Kratos Integral Systems International, США, разработан язык управления КА [12], применяющийся в составе СПО-У Ares. Скрипт представляет собой текстовый файл в виде таблицы с использованием символов табуляции в качестве разделителей со следующими столбцами: номер шага, инструкция, команда, ТМ-параметр, метка, вызов функции. Позволяет вызывать функции из предоставляемой библиотеки функций. Помимо общих функций языка управления, дополнительно поддерживается загрузка файлов с аргументами команд.

Язык описания типовых работ — ЯОТР — разработан на базе акционерного общества «Информационные спутниковые системы» им. академика М. Ф. Решетнёва» (АО «ИСС»), Россия.

В 1983 г. на базе Института прикладной математики им. Келдыша совместно с РКК «Энергия» был разработан проблемно-ориентированный табличный язык «Диполь» для проведения наземных испытаний проекта «Энергия-Буран». В 1987 г. язык «Диполь-4» был внедрен в Научно-производственное объединение прикладной механики им. академика М. Ф. Решетнёва для написания программ наземных испытаний КА различного назначения и в 1993 г. был изменен и доработан до версии «Диполь-5». Затем, в 1999 г. на основе языка «Диполь-5» был разработан ЯОТР путем исключения команд, не применимых для управления КА; табличная структура языка, а также написание отдельных операторов было сохранено. Язык применяется для управления и эксплуатации КА разработки АО «ИСС». В 2004 г. на базе языка «Диполь-5» была разработана очередная версия языка «Диполь-6» [13], которая применяется для комплексных электрических испытаний КА производства АО «ИСС».

Язык описания типовых работ, применяемый для управления КА, — это интерпретируемый скриптовый процедурный язык с текстовой нотацией. Типовая работа (ТР) представляет собой текстовый файл в виде таблицы с использованием символов псевдографики “|” со следующими столбцами: номер строки, тип строки, оператор языка, аргументы. Пример ТР управления КА показан на рис. 1.

Результаты сопоставления перечисленных языков управления КА по общим характеристикам языков программирования (тип языка, используемая парадигма, возможность параллельного выполнения сценариев, используемая нотация) представлены в табл. 1. Анализ полученных данных показывает, что несмотря на сходные функциональные задачи, связанные с выдачей команд на КА, реализация языков управления КА сильно различается:

— языки управления могут быть как общего назначения, так и специализированными (разра-

○		1				ПРОГРАМ							
○		2				ВЫЗВАТЬ				_PK			
○		3				40		&1					
○		4				ЕСЛИТО							
		5				#ГТС_БЛ				0			
○		6				ПЕЧАТЬ							
Ф		7		Счетчик фраз		обнулен							
○		8				ИНАЧЕ							
○		9				ПЕЧАТЬ							
Ф		10		Счетчик фраз		не был изменен							
○		11				КЕСЛИТО							
○		12				КПРОГРАМ							

■ Рис. 1. Пример ТР на ЯОТР

■ Таблица 1. Сопоставление реализации ряда языков управления КА

Критерий сопоставления	Python	Pluto	SCL	Без названия, СПО-У SMACS	Без названия, СПО-У Ares	ЯОТР
Страна разработки СПО-У	США, Испания	Англия, Германия, Бразилия	США	Япония	США	Россия
Тип языка	Общего назначения	Специализированный	Специализированный	Специализированный	Специализированный	Специализированный
Используемая парадигма	Объектно-ориентированный	Процедурный	Процедурный	Процедурный	Процедурный	Процедурный
Поддержка параллельного выполнения скриптов управления	+	–	+	–	–	–
Табличная/текстовая нотация	Текстовая	Текстовая	Текстовая	Табличная	Текстовая в виде таблицы с использованием символов псевдографики	Текстовая в виде таблицы с использованием символов псевдографики

ботанными специально для автоматизации процесса управления КА);

— возможно использование как процедурно-го, так и объектно-ориентированного стилей программирования;

— отдельные языки поддерживают только последовательное выполнение сценариев управления, другие — как последовательное, так и параллельное;

— для написания сценариев может применяться как текстовое, так и табличное представление.

Требования к создаваемому языку

Многолетний опыт эксплуатации ЯОТР в качестве языка управления КА выявил в нем ряд ограничений.

1. Использование текстово-табличной нотации крайне затруднительно, так как она требует запоминания правил заполнения каждого отдельного поля и различается от одного оператора к другому, что особенно неудобно, например, в условиях, когда необходимо создать ТР за ограниченный промежуток времени (для обработки нештатной ситуации в ходе испытаний или управления КА).

2. Не поддерживается английский язык написания ТР, что является несоблюдением одного из требований международных контрактов, заключаемых АО «ИСС».

3. Отсутствуют возможности для автоматизации выполнения отдельных функций, необходимых для проведения процесса управления современными КА.

Последний недостаток следует уточнить:

— невозможно использовать номер КА в качестве параметра, по значению которого меняется логика ТР для систем управления группировками КА;

— для ситуации множественного выбора не предусмотрена обработка случая, когда ни один из предложенных вариантов не подошел;

— не поддерживается использование ТМ-параметров в качестве аргументов операций (сравнение, ожидание, выбор из нескольких вариантов и т. д.);

— отсутствует обработка ТМ-параметров из заранее сохраненного архива (например, для восстановления состояния бортового ретрансляционного комплекса в исходное состояние после какого-либо сбоя);

— невозможно использовать дробные числа в качестве параметров для проверки различных условий (например, принадлежности ТМ-параметра некоторому интервалу и т. д.);

— не поддерживается возможность выдачи команды смены декодера средствами языка.

Применение зарубежных аналогов в качестве языка управления является нецелесообразным из-за проблем интеграции с существующими средствами управления КА, отсутствия детального описания языков в открытом доступе и пр. Это привело к необходимости разработки нового проблемно-ориентированного языка управления КА.

Для соответствия набору предъявляемых требований разрабатываемый язык должен обладать следующими характеристиками:

— быть императивным (описывать логику программы в виде последовательности инструкций);

— иметь простой синтаксис (т. е. содержать минимум синтаксических конструкций, что упрощает процесс обучения конечных пользователей, увеличивает скорость разработки сценариев, и вследствие простоты получаемых программ управления уменьшает количество потенциальных ошибок);

— использовать термины, близкие к предметной области, с сохранением преемственности с существующим языком управления, что упрощает восприятие нового языка пользователями;

— содержать как русский, так и английский варианты написания операторов языка для использования в отечественных и международных проектах АО «ИСС».

Язык управления КА должен поддерживать следующие функции:

— выдачу управляющих воздействий, например команд управления КА и массивов командно-программной информации;

— проверку выполнения заданного условия, например попадания текущего значения указанного ТМ-параметра в допустимый интервал;

— проверку многовариантного ветвления, включая случай, когда ни один вариант из предложенных не подошел;

— проверку выполнения нескольких заданных логических условий «И», «ИЛИ», «НЕ»;

— задание паузы на указанный промежуток времени;

— печать текстовых сообщений оператору;

— комментарии, содержащие пояснения к тексту программы;

— вызов других ТР из данной с возможностью передачи некоторых параметров, что обеспечивает повторное использование уже созданных и отлаженных сценариев, например путем вызова на выполнение одного сценария из другого;

— остановку выполнения ТР, что применяется для обработки нештатной ситуации в управлении, прерывания стандартного хода выполнения типовой процедуры в случае, когда необходимо вмешательство оператора для анализа сложившейся ситуации;

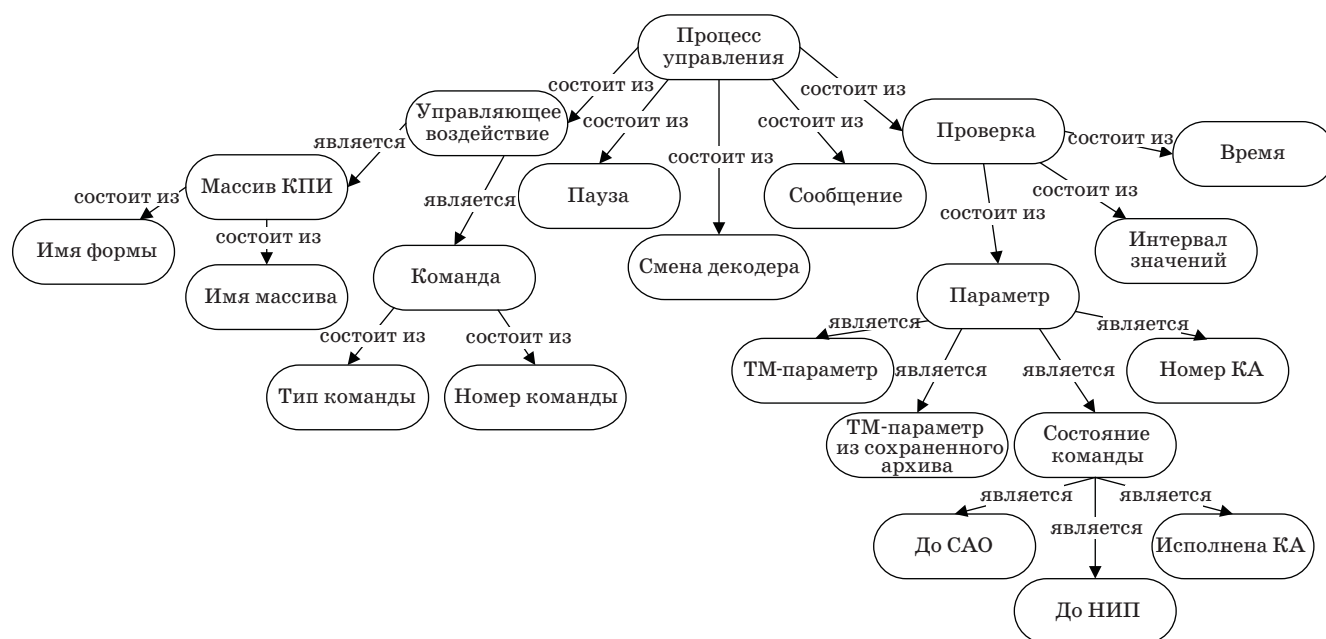
— циклы, необходимые для выполнения повторяющихся действий.

Разработка языка управления КА

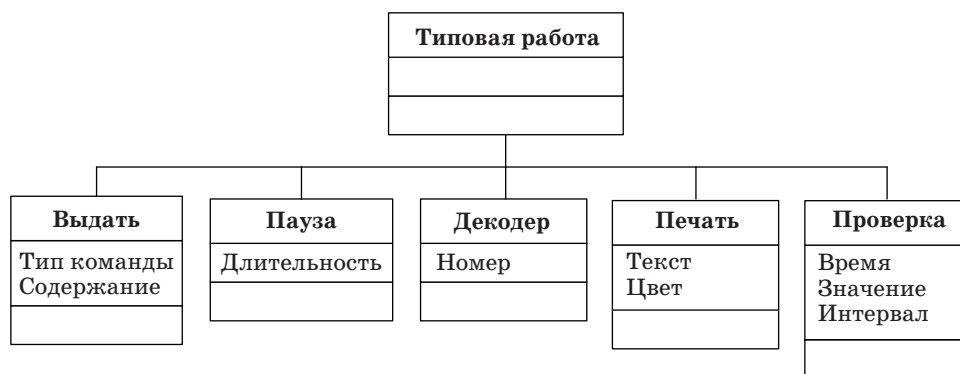
Рассмотренные выше требования были реализованы в языке управления КА «Дельта», который относится к классу предметно-ориентированных языков.

В соответствии с работой [14] последовательность разработки внешнего проблемно-ориентированного языка состоит из трех шагов: определение семантической модели; определение синтаксической модели (абстрактный и конкретный синтаксис); определение правил трансформации (правила, по которым абстрактное представление транслируется в исполнимое). Третий этап подразумевает разработку транслятора языка, описание которого выходит за рамки данной статьи.

Семантическую модель предметной области разрабатываемого языка представим в виде семантической сети (рис. 2).



■ Рис. 2. Семантическая модель предметной области: КПИ — командно-программная информация; CAO — система автоматизированного обмена; НИП — наземный измерительный пункт



■ Рис. 3. Мета модель разрабатываемого языка «Дельта»

Абстрактный синтаксис — описание синтаксиса, не зависящее от способа представления данных (от текстового представления команд или пиктограмм, используемых для визуализации объектов). Мета модель разрабатываемого языка, созданная на основе анализа семантической модели предметной области, представлена с помощью диаграммы классов UML на рис. 3.

Данная мета модель несколько отличается от существующих универсальных метамоделей языка управления КА [5, 12], что объясняется отличиями в автоматизируемых технологиях управления.

Конкретный синтаксис — набор правил, определяющих допустимые цепочки входных символов. Из возможных форм представления (текстовой, табличной, графической) выбрано текстовое представление. Это связано с тем, что автоматизируемая логика управления КА является линейной, т. е. относительно простой для восприятия в виде текста.

Общие характеристики

Сценарий является текстовым файлом, хранящимся в базе данных и обладающим уникальным именем, по которому его можно вызвать на исполнение в ходе сеанса управления. Также его можно вызывать из других сценариев, для чего в состав сценария могут входить формальные параметры, которые при вызове должны быть заменены на фактические. В случае отсутствия такой замены выдается сообщение об ошибке.

Операторы сценария выполняются последовательно, исключая случаи обработки управляющих конструкций (операторов ветвления).

Переменные

— Имена ТМ-параметров (идентификаторы которых представляют собой произвольную последовательность символов и могут меняться от одной космической системы к другой). Текущие значения ТМ-параметров — числа с плавающей точкой.

— ТМ-параметр из сохраненного архива. При необходимости оператор может сохранить текущие значения ТМ-параметров на определенный момент времени в некоторый архив. Для использования сохраненных значений необходимо указать их в формате «#имя_тм_параметра».

— Состояние выданной команды (булева переменная) — на каком этапе прохождения она находится: дошла до программного обеспечения САО, куда выданные команды попадают перед тем, как быть отправленными на наземную станцию («#доСАО»), выдана наземной станцией («#доНИП»), выполнена на КА («#ИСП»).

— Номер КА — применяется для систем из нескольких КА, для управления которых применима одна и та же ТР. Номер КА должен быть указан в формате «#КА».

Операторы

Перечень операторов языка «Дельта» создан на основе разработанной мета модели (см. рис. 3). Многие ключевые слова, используемые в языке, совпадают с терминами предметной области.

Оператор выдачи управляющих воздействий на КА. В качестве аргументов оператора должны быть использованы тип и номер команды. Допускаются следующие типы команд: разовые команды (РК), программные команды (ПК) и некоторые другие, зависящие от конкретной реализации космической системы. Помимо этого можно выдать массив КПИ, также обладающий уникальным именем и хранящийся в специализированной таблице баз данных. Применяемое ключевое слово для обозначения оператора — «ВЫДАТЬ»:

```
<оператор _ выдать> ::= ВЫДАТЬ <тип _ команды>
<номер _ команды>
```

Пример использования:

```
ВЫДАТЬ РК 40
```

Оператор выдачи произвольных сообщений на экран. Ключевое слово для обозначения оператора — «ПЕЧАТЬ». Цвет выводимого сообщения можно указать, включив в сообщение

наименование цвета в формате «#ЦВЕТ», например «#КРАСНЫЙ» или «#ЗЕЛЕНый»:

```
<оператор _ печать> ::= ПЕЧАТЬ (<цвет>) <сообщение>
<цвет> ::= #КРАСНЫЙ | #ЗЕЛЕНый | #СИНИЙ |
#ОРАНЖЕВый | #РОЗОВый | #ФИОЛЕТОВый |
#ЖЕЛТый
```

Пример использования:

```
ПЕЧАТЬ Выполнение ТР остановлено #КРАСНый.
```

Оператор остановки выполнения программы на указанный интервал времени. Ключевое слово — «ПАУЗА»:

```
<оператор _ пауза> ::= ПАУЗА <время>
<время> ::= <чч>: <мм>: <сс>
```

Пример использования:

```
ПАУЗА 00:01:00
```

Оператор принудительной остановки процесса выполнения ТР. Ключевое слово — «СТОП». При необходимости выполнение ТР можно продолжить СПО-У:

```
<опер _ стоп> ::= СТОП
```

Пример использования:

```
СТОП
```

Оператор смены декодера. Ключевое слово — «ДЕКОДЕР», затем идет номер активируемого декодера, «А» или «В»:

```
<опер _ декодер> ::= ДЕКОДЕР <номер _ декодера>
```

Пример использования:

```
ДЕКОДЕР А
```

Помимо целевых операторов, описывающих действия по управлению КА, в язык должны быть введены управляющие конструкции, обеспечивающие возможность нелинейного выполнения скрипта управления.

Условные операторы. Операции «ЕСЛИ», «ИНАЧЕ», «КЕСЛИ» позволяют выполнять определенную последовательность действий в зависимости от выполнения некоторого условия. В случае если условие выполнилось, исполняется указанная далее последовательность операций. Внутри операции «ЕСЛИ» допускается использовать оператор «ИНАЧЕ», который задает альтернативную последовательность действий на тот случай, когда условие, заданное в дополнительных строках оператора «ЕСЛИ», не выполнено. В любом случае оператор «ЕСЛИ» должен быть закрыт оператором «КЕСЛИ»:

```
<оператор _ если> ::=
ЕСЛИ(<список _ условий> ЗА <время>
<последовательность _ операторов>
ИНАЧЕ
<последовательность _ операторов>
КЕСЛИ
<время> ::= <чч>: <мм>: <сс>
<список _ условий> ::= <условие> (И <список _
условий> | ИЛИ <список _ условий>
<условие> ::= (НЕ) <параметр> <оператор> <значение> | <параметр> В <интервал>
<оператор> ::= < | > | <= | >=
```

```
<интервал> ::= (<значение>; <значение>) |
[<значение>; <значение>] | [<значение>; <значение>]
| (<значение>; <значение>]
```

Пример использования:

```
ЕСЛИ(ТМ1>8 ЗА 00:01:00)
```

```
ПЕЧАТЬ ТМ1 > 8
```

```
ИНАЧЕ
```

```
СТОП
```

```
КЕСЛИ
```

Конструкции выбора. Операторы «ВЫБОР», «ВАРИАНТ», «ОСТАЛЬНОЕ» «КВЫБОР» образуют конструкцию многовариантного ветвления в ТР — аналог оператора «SWITCH» в языках высокого уровня. Конкретная операция «ВАРИАНТ» выбирается для исполнения в том случае, если текущее проверяемое значение переменной оказалось в указанном диапазоне. Внутри операции «ВЫБОР» может содержаться операция «ОСТАЛЬНОЕ», которая будет выполняться, если ни один из предложенных вариантов не исполнился.

```
<оператор _ выбор> ::= ВЫБОР (<параметр>)
```

```
<список _ вариантов>
```

```
КВЫБОР
```

```
<список _ вариантов> ::=
```

```
ВАРИАНТ(<значение>|<интервал>)
```

```
<последовательность _ операторов>|<список _
вариантов>
```

```
<интервал> ::= (<значение>; <значение>) |
[<значение>; <значение>] | [<значение>; <значение>]
| (<значение>; <значение>]
```

Пример использования:

```
ВЫБОР (ТМ1)
```

```
ВАРИАНТ (.., 8)
```

```
ПЕЧАТЬ Параметр ТМ1 < 8
```

```
ВАРИАНТ 8
```

```
ПЕЧАТЬ Параметр ТМ1 = 8
```

```
ВАРИАНТ (8, ..)
```

```
ПЕЧАТЬ Параметр ТМ1 > 8
```

```
КВЫБОР
```

Циклы. Операция обеспечивает повторение выбранных операций по управлению указанное количество раз:

```
<оператор _ цикл> ::= ЦИКЛ (<число раз>)
```

```
<последовательность _ операторов>
```

```
КЦИКЛ
```

Пример использования:

```
ЦИКЛ (8)
```

```
ВЫДАТЬ РК 40
```

```
ПАУЗА 00:01:00
```

```
КЦИКЛ
```

Конструкции включения. Оператор «ВЫЗВАТЬ» инициирует выполнение указанной ТР из текущей. При необходимости можно передать в вызываемую ТР значения некоторых параметров. В тексте вызываемой ТР должны быть указаны параметры для замены в формате «&1», «&2», которые будут заменены переданными при вызове

значениями. В случае отсутствия замены параметра ситуация считается ошибочной, о чем выдается соответствующее сообщение:

```
<опер _ вызвать>::= ВЫЗВАТЬ <имя _ работы>
(<список _ параметров>)
<список _ параметров> ::= <значение>(<список _
параметров>)
```

Пример использования:

```
ВЫЗВАТЬ _ РК(40)
```

Комментарии. Должна быть предусмотрена возможность включения в текст процедуры произвольных пояснений к особенностям ее написания, не обрабатываемых интерпретатором языка. Для включения комментария должно использоваться ключевое слово «К»:

```
<опер _ комментарий>::= К <текст _ комментария>
```

Пример использования:

```
К Дата создания ТР – 2015.04.29
```

Для международных проектов АО «ИСС» поддерживается также вариант на английском языке, сохраняющий все особенности разработанного синтаксиса, за исключением написания ключевых слов. Соответствие русскоязычных и англоязычных слов представлено в табл. 2.

■ **Таблица 2.** Соответствие ключевых слов в русском и английском вариантах разработанного языка управления

Русский вариант	Английский вариант
ВЫДАТЬ	SEND
ЕСЛИ	IF
ЗА	IN
КЕСЛИ	EIF
ВЫБОР	SELECT
ВАРИАНТ	CASE
ОСТАЛЬНОЕ	REST
КВЫБОР	ESELECT
И	AND
ИЛИ	OR
НЕ	NOT
ПАУЗА	PAUSE
ПЕЧАТЬ	PRINT
К	K
ВЫЗВАТЬ	CALL
ДЕКОДЕР	DECODER
СТОП	STOP
ЦИКЛ	CYCLE
КЦИКЛ	ECYCLE
#НКА	#SC
#доСАО	#toFES
#доЗС	#toST
#ИСП	#DONE
#КРАСНЫЙ	#RED
#ЗЕЛЕНый	#GREEN
#СИНИЙ	#BLUE
#ОРАНЖЕВый	#ORANGE
#РОЗОВый	#PINK
#ФИОЛЕТОВый	#MAGENTA
#ЖЕЛТый	#YELLOW

К Пример написания типовой работы

```
ВЫЗВАТЬ РТР _Вкл_1
ВЫЗВАТЬ РТР _АКТ(40, ф 3001)
ПАУЗА 00:00:10
```

```
ЕСЛИ (АТМ 1 = 1 ЗА 00:00:20)
    ПЕЧАТЬ #КРАСНЫЙ Параметр АТМ 1 = 1
    СТОП
КЕСЛИ
```

```
ПАУЗА 00:00:10
```

```
ЕСЛИ (АТМ 1 = 1 И АТМ 2 В (-5,1) И АТМ3 В [-5,1] ЗА 00:00:20)
    ПЕЧАТЬ ТМ параметры в норме
    СТОП
КЕСЛИ
```

```
ДЕКОДЕР А
```

```
ЦИКЛ (8)
    ВЫДАТЬ РК 40
    ПАУЗА 00:01:00
КЦИКЛ
```

```
ЕСЛИ (#ИСП = 0)
    ПЕЧАТЬ #КРАСНЫЙ Команда не прошла !
    СТОП
КЕСЛИ
```

```
ВЫБОР АТМ2
ВАРИАНТ (-5,1)
    ПЕЧАТЬ Параметр АТМ2 находится в границах(-5,1)
ВАРИАНТ (1,8)
    ПЕЧАТЬ Параметр АТМ2 находится в границах(1,8)
ОСТАЛЬНОЕ
    ПЕЧАТЬ #КРАСНЫЙ Параметр АТМ2
                                не принял требуемых значений
    СТОП
КВЫБОР
```

■ **Рис. 4.** Пример ТР на языке управления «Дельта»

Пример ТР на разработанном языке управления «Дельта» показан на рис. 4.

Стоит отметить, что разработанный язык «Дельта» поддерживает и расширяет функциональные возможности ранее эксплуатируемого ЯОТР для обеспечения переносимости управления в уже существующих системах управления КА.

Закключение

В ходе выполнения исследования были рассмотрены существующие зарубежные аналоги — языки управления космическими аппаратами, выявлена схожесть выполняемых ими функциональных задач и отсутствие общих тенденций в реализации.

Проанализирован эксплуатируемый на АО «Информационные спутниковые системы» язык управления космическими аппаратами — ЯОТР, сформулированы его недостатки. Выявлены требования к новому языку управления КА: импера-

тивность; простой синтаксис; использование терминов, близких предметной области; наличие как русского, так и английского вариантов написания операторов языка; необходимость в расширении существующих функциональных возможностей.

В соответствии со сформулированными требованиями был разработан новый язык управления КА «Дельта», удовлетворяющий текущим функциональным задачам предметной области и обеспечивающий лаконичность и простоту соз-

даваемых сценариев управления. Язык является более простым в освоении для специалистов по управлению КА, чем предыдущий вариант языка управления, что увеличивает скорость разработки сценариев управления и уменьшает количество потенциальных ошибок.

Данная разработка предназначена для управления большинством КА производства АО «ИСС» как в отечественных, так и в международных проектах.

Литература

1. Соловьев В. А., Лысенко Л. Н., Любинский В. Е. Управление космическими полетами. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2010. — 432 с.
2. Иванов Д. С. Управление полетом космического аппарата. <http://www.cosmos-journal.ru/articles/936/> (дата обращения: 03.06.2015).
3. Chaudhri G., Carter J., Kizzort B. A Model for a Spacecraft Operations Language. <http://arc.aiaa.org/doi/pdf/10.2514/6.2006-5708> (дата обращения: 03.06.2015)
4. ECSS-E-ST-70-32C Space Engineering Test and Operations Procedure Language. [ftp://ftp.heanet.ie/disk1/sourceforge/o/op/open-hades/Documentation/ECSS-E-ST-70-32C%20\(31July2008\).pdf](ftp://ftp.heanet.ie/disk1/sourceforge/o/op/open-hades/Documentation/ECSS-E-ST-70-32C%20(31July2008).pdf) (дата обращения: 03.06.2015).
5. Satellite Operations Language Metamodel (SOLM). <http://www.omg.org/spec/SOLM/1.0/PDF/> (дата обращения: 03.06.2015).
6. Hauck T. F. Use of the Ground Support Equipment Operating System (GSEOS) Software on the Messenger Mission: a Case Study. <http://descanso.jpl.nasa.gov/RCSGSO/Paper/A0013Paper.pdf> (дата обращения: 03.06.2015).
7. Garcia G., et al. Migration of the Operations of a Legacy Multi-Satellite System to a State-of-the-art Ground System. <http://www.enu.kz/repository/2009/AIAA-2009-6622.pdf> (дата обращения: 03.06.2015).
8. Seymour M. A. The PLUTO Operations Procedure Language and its Use for RADARSAT-2 Mission Operations. <http://arc.aiaa.org/doi/pdf/10.2514/6.2004-538-333> (дата обращения: 03.06.2015).
9. Mims T. L. Use of Spacecraft Command Language for Advanced Command and Control Applications. <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110015644.pdf> (дата обращения: 03.06.2015).
10. Flight Software — SCL for Flight Systems. <http://www.sra.com/scl/flight/> (дата обращения: 03.06.2015).
11. Konishi M., Miyoshi M., Sugawara M. Spacecraft Control System Development. <http://arc.aiaa.org/doi/pdf/10.2514/6.2002-T3-27> (дата обращения: 03.06.2015).
12. Cruickshank D. M., Aparicio A. J., Mistry A. A., Poynton B. R. Using ARES to Improve Efficiency in Satellite Operations. http://www.powershow.com/view/1c44e-NjEwM/Using_ARES_to_Improve_Efficiency_in_Satellite_Operations_powerpoint_ppt_presentation (дата обращения: 03.06.2015).
13. Барков А. В. Структура проблемно-ориентированного языка испытаний космических аппаратов // Вестник Сибирского государственного аэрокосмического университета им. академика М. Ф. Решетнёва. 2006. № 5. С. 15–18.
14. Ботов Д. С. Обзор современных средств создания и поддержки предметно-ориентированных языков программирования // Вестник Южно-Уральского государственного университета. 2013. № 1. С. 10–15.

UDC 004.434

doi:10.15217/issn1684-8853.2015.5.82

Spacecraft Control Language Development

Kosmynina N. A.^a, Engineer, kosmynina@iss-reshetnev.ru

Legalov A. I.^b, Dr. Sc., Tech., Professor, legalov@mail.ru

Barkov A. V.^a, PhD, Tech., Head of Section, barkov@iss-reshetnev.ru

Lapin A. A.^a, PhD, Tech., Head of Section, lapin@iss-reshetnev.ru

^aJoint-stock Company «Academician M. F. Reshetnev «Information Satellite Systems», 52, Lenin St., 662972, Zheleznogorsk, Krasnoyarsk region, Russian Federation

^bSiberian Federal University, 79, Svobodny Pr., 660041, Krasnoyarsk, Russian Federation

Purpose: Nowadays, the capabilities of the control language used in “JSC Academician M.F. Reshetnev Information Satellite Systems” do not meet the requirements of the object domain: the language does not support new functional tasks, and it is too sophisticated

to learn it quickly. A new control language needs to be developed. **Results:** We have analyzed the limitations of the current control language, the object domain requirements and the existing analogs. On the base of this analysis, a new control language named “Delta” was designed and developed. It has simple syntax, using the object domain terms as key words and supporting both English and Russian syntax in spelling its operators. “Delta” supports basic control functions: sending commands, incoming telemetry analysis, pauses, messages, conditional operators, etc. **Practical relevance:** “Delta” is intended to be used to control the majority of “JSC Academician M. F. Reshetnev Information Satellite Systems” satellites, in both domestic and foreign projects.

Keywords — Problem-Oriented Language, Domain Specific Language, Satellite Control, Control Script.

References

1. Solov'ev V. A., Lysenko L. N., Liubinskii V. E. *Upravlenie kosmicheskimi poletami* [Satellites Flight Control]. Moscow, MGTU im. N. E. Bauman Publ., 2010. 432 p. (In Russian).
 2. Ivanov D. S. *Upravlenie poletom kosmicheskogo apparata* [Spacecraft Flight Control]. Available at: <http://www.cosmos-journal.ru/articles/936/> (accessed 03 June 2015).
 3. Chaudhri G., Carter J., Kizzort B. *A Model for a Spacecraft Operations Language*. Available at: <http://arc.aiaa.org/doi/pdf/10.2514/6.2006-5708> (accessed 03 June 2015).
 4. ECSS-E-ST-70-32C *Space Engineering Test and Operations Procedure Language*. Available at: [ftp://ftp.heanet.ie/disk1/sourceforge/o/op/open-hades/Documentation/ECSS-E-ST-70-32C%20\(31July2008\).pdf](ftp://ftp.heanet.ie/disk1/sourceforge/o/op/open-hades/Documentation/ECSS-E-ST-70-32C%20(31July2008).pdf) (accessed 03 June 2015).
 5. *Satellite Operations Language Metamodel (SOLM)*. Available at: <http://www.omg.org/spec/SOLM/1.0/PDF/> (accessed 03 June 2015).
 6. Hauck T. F. *Use of the Ground Support equipment Operating System (GSEOS) Software on the Messenger Mission: a Case Study*. Available at: <http://descanso.jpl.nasa.gov/RCSGO/Paper/A0013Paper.pdf> (accessed 03 June 2015).
 7. Garcia G., Palsson M., Morel T., Beech T., Honold P., Fuente O., Cervino A., Cunha C., Florido L., Hilden V., Silva R. *Migration of the Operations of a Legacy Multi-Satellite System to a State-of-the-art Ground System*. <http://www.enu.kz/repository/2009/AIAA-2009-6622.pdf> (accessed 03 June 2015).
 8. Seymour A. M. *The PLUTO Operations Procedure Language and its Use for RADARSAT-2 Mission Operations*. Available at: <http://arc.aiaa.org/doi/pdf/10.2514/6.2004-538-333> (accessed 03 June 2015).
 9. Mims T. L. *Use of Spacecraft Command Language for Advanced Command and Control Applications*. Available at: <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110015644.pdf> (accessed 03 June 2015).
 10. *Flight Software – SCL for Flight Systems*. Available at: <http://www.sra.com/scl/flight/> (accessed 03 June 2015).
 11. Konishi M., Miyoshi M., Sugawara M. *Spacecraft Control System Development*. <http://arc.aiaa.org/doi/pdf/10.2514/6.2002-T3-27> (accessed 03 June 2015).
 12. Cruickshank D. M., Aparicio A. J., Mistry A. A., Poynton B. R. *Using ARES to Improve Efficiency in Satellite Operations*. http://www.powershow.com/view/1c44e-NjEwM/Using_ARES_to_Improve_Efficiency_in_Satellite_Operations_powerpoint_ppt_presentation (accessed 03 June 2015).
 13. Barkov A. V. Structure of the Problem-oriented Language of Spacecrafts Tests. *Vestnik Sibirskogo gosudarstvennogo aerokosmicheskogo universiteta im. akademika M. F. Reshetneva*, 2006, no. 5, pp. 15–18 (In Russian).
 14. Botov D. S. Review of Modern Development and Support Tools for Domain-specific Programming Language. *Vestnik Iuzhno-Ural'skogo gosudarstvennogo universiteta*, 2013, no. 1, pp. 10–15 (In Russian).
-

УВАЖАЕМЫЕ АВТОРЫ!

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

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

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

АРХИТЕКТУРА ДЕЦЕНТРАЛИЗОВАННОЙ РЕКОМЕНДУЮЩЕЙ СИСТЕМЫ, ОСНОВАННОЙ НА ПРИМЕНЕНИИ ЛОКАЛЬНО-ЧУВСТВИТЕЛЬНОГО ХЕШИРОВАНИЯ

А. В. Пономарев^а, канд. техн. наук, старший научный сотрудник

^аСанкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург, РФ

Постановка проблемы: рекомендующие системы широко используются в современных системах электронной коммерции, помогая пользователям ориентироваться в многообразии предлагаемых товаров и услуг. Наибольшее распространение получили централизованные архитектуры построения таких систем. Однако централизация влечет за собой ряд недостатков, среди которых — необходимость передачи пользователем сведений о предпочтениях стороне, осуществляющей эксплуатацию такой системы, и наличие единой точки отказа. **Цель:** построение децентрализованной рекомендующей системы, в которой для формирования рекомендаций используется сходство предпочтений пользователей (коллаборативная фильтрация), но полные сведения о предпочтениях хранятся только на узле, контролируемом самим пользователем, и не передаются другим узлам. **Результаты:** предложена архитектура децентрализованной рекомендующей системы, включающая структурированную одноранговую сеть, в которой каждый узел соответствует одному пользователю и хранит профиль его предпочтений, и специальный узел для информационного согласования участников сети. В качестве механизма, обеспечивающего, с одной стороны, поиск пользователей со схожими предпочтениями, а с другой стороны, ограниченное раскрытие информации о предпочтениях, используется локально-чувствительное хеширование. Для повышения уровня приватности пользователей в одноранговой сети применяется схема анонимизации. **Практическая значимость:** предложенный подход является достаточно универсальным и может быть использован для построения систем коллаборативной фильтрации в различных прикладных областях.

Ключевые слова — локально-чувствительное хеширование, одноранговые сети, рекомендующие системы, коллаборативная фильтрация.

Введение

Большинство широко распространенных подходов к построению рекомендующих систем предполагают централизованную архитектуру. Важным достоинством централизации является существование широкого спектра техник моделирования предпочтений пользователей, предполагающих доступ к профилям всех пользователей (большинство реализаций метода ближайших соседей, разложение матрицы предпочтений и др.). Кроме того, при централизованном хранении информации о предпочтениях сторона, осуществляющая эксплуатацию рекомендующей системы, может производить всевозможные исследования этой информации, в том числе и не связанные напрямую с формированием рекомендаций.

Однако централизованная архитектура не свободна и от недостатков. Во-первых, в централизованных рекомендующих системах естественным образом возникает неоднозначная ситуация, касающаяся прав на информацию о предпочтениях. Как правило, пользователь не знает, какая именно информация о его действиях собирается, и не может извлечь (или уничтожить) эту информацию из системы. Более того, в случае прекращения функционирования сервиса, включавшего такую рекомендующую систему, соответ-

ствующая информация может быть безвозвратно утеряна. Во-вторых, централизация влечет за собой определенное разделение профиля пользователя. Пользователь может взаимодействовать с несколькими рекомендующими системами, предоставляя каждой лишь некоторые аспекты своих предпочтений. В результате предпочтения оказываются распределены между этими системами, хотя их консолидация могла бы улучшить качество рекомендаций. Наконец, централизация приводит к уменьшению надежности системы в целом за счет появления единой точки отказа, хотя в современных компьютерных системах этот недостаток в значительной мере ослабляется многоуровневыми схемами дублирования и репликации.

Децентрализация рекомендующей системы позволяет добиться двух важных целей:

- распределения функции формирования рекомендаций между пользователями и, как следствие, снятия необходимости в дорогостоящем сервере и повышения масштабируемости системы;
- повышения уровня приватности пользователей, поскольку исчезает необходимость в передаче предпочтений центральному серверу.

Есть несколько подходов к децентрализации рекомендующих систем. В этой статье развивается подход, в соответствии с которым пользователь хранит все сведения о предпочтениях только

на своем компьютере. Это полностью снимает упомянутую выше неоднозначную ситуацию, касающуюся прав на информацию о предпочтениях. Это также снимает проблему разделения профиля пользователя, поскольку все предпочтения сосредоточиваются на одном устройстве, контролируемом пользователем. При необходимости получения рекомендаций устройство посылает запросы на предоставление рекомендаций устройствам других пользователей.

И хотя в данном подходе устраняются все перечисленные недостатки централизованных рекомендующих систем, возникает и ряд трудностей. Главная проблема — ее решению посвящена и эта статья — состоит в реализации рекомендующего алгоритма, не требующего от пользователя передачи профиля своих предпочтений третьим лицам (участникам распределенной сети рекомендаций). Здесь следует сделать небольшое уточнение. Существует два основных класса рекомендующих систем: контентные системы и системы коллаборативной фильтрации. В контентных системах для формирования рекомендаций используются свойства самих объектов — система рекомендует объекты, похожие (с точки зрения некоторого формального представления) на те, что были полезны пользователю в прошлом. В системах же коллаборативной фильтрации сами свойства объектов не анализируются, система рекомендует те объекты, которые были высоко оценены пользователями, демонстрирующими схожие предпочтения. Конечно, трудности, связанные с децентрализацией, преимущественно касаются систем коллаборативной фильтрации, в основе которых лежит анализ сходства между пользователями, сопоставление их предпочтений, которое и осложняется распределенной организацией системы. Речь далее пойдет именно о таких системах, и под рекомендующей системой будет, если не оговорено иное, пониматься частный случай — система коллаборативной фильтрации.

В данной статье предложена архитектура рекомендующей системы, включающая структурированную одноранговую (P2P) сеть, в которой каждый узел соответствует одному пользователю и хранит профиль его предпочтений, и специальный узел для информационного согласования участников сети. Такой подход может быть классифицирован как гибридная одноранговая сеть, в которой часть функций выполняется исключительно посредством взаимодействия между равноправными узлами, а часть функций требует наличия специального узла. Предлагаемая архитектура обеспечивает ограниченное раскрытие предпочтений — не существует способа одновременно получить оценки, которые пользователь присвоил объектам, и сетевой адрес пользователя без глобального контроля над сетью.

Сама по себе задача построения рекомендующих систем, основанных на одноранговых сетях, не является новой. Существует определенный пласт работ, в которых эта задача ставится и предлагаются различные подходы к ее решению.

В системе P2Pres [1, 2] для распространения запросов и рекомендаций используется комбинация так называемого «дружеского» подхода к построению структуры сети (*friend-of-a-friend*), когда связи устанавливаются только между знакомыми пользователями, и лавинных алгоритмов распространения запросов.

В ряде описанных методов происходит построение оверлейной структуры, соответствующей близости интересов пользователей, поверх одноранговой сети [3, 4]. Рекомендации формируются поиском по оверлейной структуре на определенную глубину. Одним из распространенных алгоритмов такого «выравнивания» структуры сети под отношения между узлами является T-Map [5]. В данной работе сами оценки пользователя не раскрываются узлом сети, поэтому напрямую использовать T-Map или какой-либо схожий алгоритм нельзя из-за невозможности определить сходство узлов.

Другой подход заключается в применении алгоритмов случайного блуждания для поиска схожих узлов [6]. Для получения рекомендаций достаточно сформировать случайную выборку узлов сети, а затем использовать ближайшие, в соответствии с заданной мерой сходства, узлы из этой выборки [7].

Есть также работы, в которых авторы исследуют возможность применения структурированных одноранговых сетей для построения рекомендующих систем. Например, в работах [8, 9] оценки, присваиваемые объектам пользователями, сохраняются в распределенной хеш-таблице (*Distributed Hash Table* — DHT). Отличие предлагаемого в данной статье подхода заключается в том, что в распределенную хеш-таблицу помещаются не оценки, а сами узлы, и механизм быстрого поиска по этой таблице используется для поиска узлов, соответствующих пользователям со схожими интересами.

Формирование рекомендаций с помощью локально-чувствительного хеширования

Локально-чувствительное хеширование (ЛЧХ) — это широко распространенный метод приближенного решения задачи поиска k ближайших соседей. Идея метода состоит в построении такой хеш-функции многомерных объектов, чтобы схожие объекты с высокой вероятностью получали одинаковое значение хеш-функции. Методы и алгоритмы поиска ближайших соседей находят широкое применение при построении

рекомендующих систем. Одним из основополагающих допущений коллаборативной фильтрации является представление о том, что пользователи, имевшие схожие предпочтения в прошлом, вероятно, имеют схожие предпочтения в настоящем, что может быть использовано при формировании рекомендаций. Если представить предпочтения пользователя в виде вектора и ввести соответствующую меру близости, то поиск пользователей со схожими интересами можно будет интерпретировать как поиск ближайших соседей.

В этом подразделе приводится формальное описание коллаборативной фильтрации, основанной на локально-чувствительном хешировании.

Идея ЛЧХ

Описание базовых идей ЛЧХ приводится в соответствии с работой [10]. Пусть $d_1 < d_2$ — два значения расстояния в соответствии с некоторой мерой d . Семейство функций F называется (d_1, d_2, p_1, p_2) -чувствительным, если для каждой функции f в F :

— если $d(a, b) \leq d_1$, то вероятность того, что $f(a) = f(b)$, не меньше p_1 ;

— если $d(a, b) \geq d_2$, то вероятность того, что $f(a) = f(b)$, не больше p_2 .

Важной идеей в теории ЛЧХ является усиление, в основе которого лежат понятия И-конструкции и ИЛИ-конструкции, определенные ниже.

Пусть задано (d_1, d_2, p_1, p_2) -чувствительное семейство функций F , новое семейство функций F' может быть получено посредством И-конструкции или ИЛИ-конструкции.

И-конструкция F' определяется следующим образом. Каждый член F' состоит из r членов F . Если f в F' и f получена из множества $\{f_1, f_2, \dots, f_r\}$ членов F , то $f(x) = f(y)$ тогда и только тогда, когда $f_i(x) = f_i(y)$ для всех $i \in \{1, \dots, r\}$. Поскольку члены F' независимо выбираются из F , F' является (d_1, d_2, p_1^r, p_2^r) -чувствительным семейством функций [10].

ИЛИ-конструкция F' определяется следующим образом. Каждый член F' состоит из b членов F . Если f в F' и f получена из множества $\{f_1, f_2, \dots, f_b\}$ членов F , то $f(x) = f(y)$ тогда и только тогда, когда существует $i \in \{1, \dots, b\}$ такой, что $f_i(x) = f_i(y)$. Аналогично F' является $(d_1, d_2, 1 - (1 - p_1)^b, 1 - (1 - p_2)^b)$ -чувствительным семейством.

Как правило, желательно, чтобы p_1 было большим, насколько возможно, а p_2 маленьким, насколько возможно. Если $p_1 < 1$, то существует вероятность того, что схожие объекты будут иметь различные значения. С другой стороны, если $p_2 > 0$, есть вероятность, что значительно различающиеся объекты получат одинаковое значение хеш-функции. Следовательно, семейство F следует выбирать таким образом, чтобы p_1 было близко

к 1, а p_2 близко к 0. Существует определенный набор хорошо изученных семейств локально-чувствительных функций, механизм усиления применяется в том случае, если только лишь средствами выбранного семейства не удастся достичь желаемых вероятностей. Если семейство F^{Ar} получено И-конструкцией r функций из семейства F , а G затем получено ИЛИ-конструкцией b функций из семейства F^{Ar} , то G является $(d_1, d_2, 1 - (1 - p_1^r)^b, 1 - (1 - p_2^r)^b)$ -чувствительным семейством. Неформально И-конструкция снижает изначально невысокую вероятность p_2 , а последующая ИЛИ-конструкция повышает изначально высокую вероятность p_1 .

Идея поиска ближайших соседей с помощью ЛЧХ описана, например, в работах [10, 11]. В первую очередь выбирается семейство F и создаются b обычных хеш-таблиц. Каждая таблица соответствует одной хеш-функции f_i^{Ar} , $i = 1, \dots, b$, где f_i^{Ar} — И-конструкция r случайных функций из F . Каждый объект x помещается в каждую из b хеш-таблиц. Ключом является значение функции $f_i^{Ar}(x)$, а значением — идентификатор объекта x или сам объект, в зависимости от задачи. При поиске ближайших соседей объекта y вычисляются $f_i^{Ar}(y)$, $i = 1, \dots, b$; все объекты, извлеченные из хеш-таблиц по полученным ключам, образуют множество кандидатов. Реальная близость оценивается уже с применением строгой меры, и происходит отсев ложно-положительных соседей из множества кандидатов.

Выбор семейства хеш-функций зависит от представления данных и функции расстояния d . Для хеммингова расстояния, например, часто применяется хеш-функция, осуществляющая выборку отдельных битов [12], для косинусной меры — метод случайных проекций [13].

В данной работе используется метод случайных проекций, т. е. функция f из семейства F соответствует одной случайной гиперплоскости; функция принимает значение 1, если хешируемая точка находится над гиперплоскостью, и 0 в противном случае.

Формирование рекомендаций

В системах коллаборативной фильтрации, основанных на сходстве пользователей (*user-based collaborative filtering*), рекомендации формируются с учетом того, в какой мере совпадают оценки пользователей, присвоенные одним и тем же объектам.

Формально, пусть r_{uj} — оценка, присвоенная объекту j пользователем u и выражающая степень того, насколько пользователю u интересен объект j или какова субъективная ценность объекта j для пользователя u . Пусть U — множество пользователей; I — множество объектов; I_u — множество объектов, оцененных пользователем u ;

I_{uv} — множество объектов, оцененных как пользователем u , так и пользователем v . Методы, основанные на сходстве, используют меру близости между пользователями, определяемую посредством сопоставления оценок, присвоенных пользователями одним и тем же объектам: $(\text{sim}(u, v) = f_s(\{r_{uj}, j \in I_{uv}\}))$, и пытаются предсказать неизвестную оценку r_{uj}^* на основе известных оценок r_{vj} и сходства между пользователями $\text{sim}(u, v)$.

В данной статье используется косинусная мера сходства между пользователями:

$$\text{sim}(u, v) = \frac{\sum_{I_{uv}} r_{uj} r_{vj}}{\sqrt{\sum_{I_{uv}} r_{uj}^2} \sqrt{\sum_{I_{uv}} r_{vj}^2}}.$$

Оценки пользователей нормализуются таким образом, что $r_{uj} = 1$ соответствует строго положительному отношению пользователя u к объекту j , а $r_{uj} = -1$ — строго отрицательному.

Рекомендующая система, использующая ЛЧХ, реализует поиск ближайших соседей. По известному набору значений хеш-функций для некоторого пользователя u система проверяет соответствующие хеш-таблицы и извлекает из них идентификаторы всех пользователей, чьи предпочтения вероятно похожи (в силу свойства хеш-функций) на предпочтения пользователя u . Затем может быть оценено точное сходство между пользователями, и объекты, высоко оцененные пользователями, похожими на u , будут рекомендованы u .

В предлагаемой системе точное значение сходства не вычисляется, поскольку это привело бы к раскрытию профиля пользователя. Вместо этого вводится приближенная мера сходства $s'(u, v)$, определяемая как количество локально-чувствительных хеш-функций, чьи значения совпали для пользователей u и v . Алгоритм рекомендации, во-первых, осуществляет поиск всех пользователей Q_u , которые могут быть соседями пользователя u , и вычисляет $s'(u, v)$ (где $v \in Q_u$). Каждому из кандидатов $v \in Q_u$ посылается запрос на список рекомендаций R_v . Предлагаемый алгоритм и система в целом предсказывают неизвестные оценки r_{uj}^* , вместо этого проводится ранжирование всех объектов, которые были рекомендованы кандидатами в соответствии с оценкой $\tilde{a}_{ui}$ объекта i для пользователя u , определяемой выражением

$$\tilde{a}_{ui} = \sum_{v \in Q_u} s'(u, v) P_{vi}^R.$$

Здесь P_{vi}^R — индикаторная функция, осуществляющая проверку того, есть ли объект i в мно-

жестве объектов, рекомендованных пользователем v :

$$P_{vi}^R = \begin{cases} 1, & i \in R_v \\ 0, & i \notin R_v \end{cases}.$$

Таким образом, в предлагаемой архитектуре профиль пользователя u представляется множеством пар (i, r_{ui}) , где i — идентификаторы объектов. Каждая из b локально-чувствительных хеш-функций представлена r векторами размерности $(|I|)$. После применения всех этих хеш-функций получается b r -мерных бинарных векторов. Полученные векторы сохраняются в хеш-таблице. Во время формирования рекомендаций производится b операций поиска в хеш-таблице, а затем каждому «кандидату», извлеченному из хеш-таблицы, посылается запрос на формирование рекомендаций. Список рекомендаций упорядочивается по значению $\tilde{a}_{ui}$.

Значения b и r являются параметрами алгоритма формирования рекомендаций. В разделе «Экспериментальное исследование» производится экспериментальная оценка того, как значения этих параметров влияют на качество рекомендаций.

Архитектура системы

Предлагаемая гибридная архитектура позволяет осуществлять обмен рекомендациями с ограниченным раскрытием предпочтений пользователя. В этом разделе описаны основные компоненты системы и сценарии их взаимодействия.

Сценарии

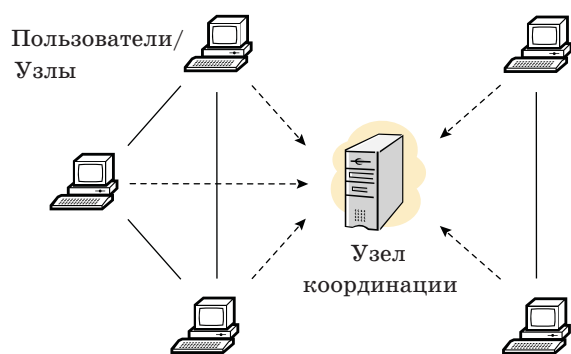
Предлагаемая система рассчитана на обеспечение двух вариантов использования: а) оценка потенциальной привлекательности объекта (или группы объектов) для данного пользователя; б) запрос рекомендаций.

Оценка потенциальной привлекательности объекта инициируется, когда необходимо проверить, может ли данный неизвестный объект быть интересен пользователю (с точки зрения логики, заложенной в систему). Пользователь передает системе идентификатор объекта, а рекомендующая система в ответ должна сообщить предполагаемую оценку привлекательности этого объекта для пользователя.

Запрос рекомендаций инициируется, когда необходимо сформировать набор новых, неизвестных пользователю объектов, которые могут оказаться ему интересны.

Компоненты

В соответствии с предлагаемым подходом рекомендующая система состоит из двух частей: одноранговой (P2P) сети рекомендаций и узла



■ Рис. 1. Связи между узлами в предлагаемой архитектуре

координации (рис. 1). Присутствие узла координации нарушает концептуальную чистоту одноранговой системы, превращая ее в гибридную, однако этот узел не играет важной роли в основных сценариях, перечисленных выше, его роль заключается в синхронизации вспомогательной информации между узлами сети.

Показаны два типа связей между узлами: связи между схожими узлами, образующими одноранговую сеть, отображены сплошными линиями; периодические связи узлов сети с узлом координации, устанавливаемые для обмена вспомогательной информацией, отображены пунктирными линиями.

1. *Одноранговая сеть рекомендаций.* В предлагаемом подходе каждый пользователь соответствует ровно одному узлу сети. На этом узле хранится вся информация о предпочтениях пользователя (в первую очередь, оценки объектов), причем узел не передает эту информацию другим узлам, он может передавать только значения локально-чувствительных хеш-функций, вычисленных от этой информации, для поиска схожих пользователей, к которым можно будет «обращаться» за получением рекомендаций.

Одноранговая сеть основана на использовании DHT [14] — распространенном подходе к построению так называемых структурированных одноранговых сетей. DHT — это класс систем, обеспечивающих хранение коллекции пар ключ — значение, распределенной по различным узлам сети, с учетом миграции фрагментов при выходе узла из состава сети.

Классические реализации подхода DHT обладают рядом уязвимостей. Для их преодоления разработано несколько анонимизированных реализаций DHT. Предлагаемая архитектура основывается на Octopus [15] — одной из таких анонимизированных реализаций. В основе таких реализаций, как правило, лежит идея построения цепочек анонимизации вместо непосредственного обращения к другому узлу сети, причем каждый узел, лежащий в такой цепочке, имеет

информацию только о соседних узлах цепочки. Таким образом, становится значительно сложнее установить, от какого же именно узла исходил запрос.

В предлагаемой системе DHT используется для хранения хеш-таблиц, применяемых в целях поиска ближайших соседей, как описано в предыдущем разделе. Каждая пара ключ — значение, хранимая в DHT, содержит информацию об одном значении локально-чувствительной хеш-функции и список узлов, соответствующих этому значению. Как уже указывалось, для поиска ближайших соседей необходимо несколько (b) хеш-таблиц. Каждая из b таблиц использует свою локально-чувствительную хеш-функцию. В данной системе предлагается хранить все b хеш-таблиц в одной DHT. Для этого ключ должен включать в себя уникальный идентификатор локально-чувствительной хеш-функции и значение этой функции.

Перед тем как включить записи в DHT, узел создает анонимизированную цепочку и использует спецификатор окончания этой цепочки как свой адрес, передаваемый другим узлам. Эти анонимизированные пути создаются при каждом очередном подключении узла к сети заново, следовательно, во время каждой новой сессии у узла оказывается новый внешний идентификатор.

Поскольку предпочтения пользователя, выраженные в оценках, присвоенных этим пользователем различным объектам, достаточно статичны, предполагается хранение ссылок на внешние «публичные» идентификаторы узлов, соответствующих пользователям со схожими интересами. Таким образом, поверх одноранговой сети образуется оверлейная сеть, сформированная ссылками между узлами пользователей со схожими интересами. Следует иметь в виду, что ссылки между вершинами в этой оверлейной сети являются не идентификаторами узлов P2P-сети, а «входами» в анонимизированные пути, ведущие к ним.

2. *Узел координации.* Распределенный характер предлагаемой системы является причиной следующей технической сложности. Для корректного вычисления локально-чувствительных хеш-функций необходимо, чтобы сами хеш-функции (т. е. гиперплоскости, которыми они представляются) были одинаковы на всех узлах. Для согласования параметров этих функций все узлы сети должны использовать один и тот же порядок следования объектов, поскольку размерность гиперплоскостей совпадает с количеством объектов и с длиной вектора пользовательских оценок. Задача поддержания глобального состояния в одноранговой сети является нетривиальной [16–18]. В предлагаемой системе для ее решения используется подход, схожий с предложенным

в статье [19] и заключающийся в отказе от чисто однорангового устройства сети. Задачей узла координации является сбор всех объектов (о которых сообщают пользователи), поддержка отношения порядка между их идентификаторами и генерация локально-чувствительных функций. Таким образом, каждый узел должен соединиться с узлом координации для двух целей: во-первых, для регистрации нового, ранее неизвестного объекта; во-вторых, для получения нового набора локально-чувствительных хеш-функций. Следует заметить, что нет необходимости генерировать новый набор хеш-функций после обнаружения каждого нового объекта. При использовании «устаревшего» набора функций получение рекомендаций оказывается возможным, но их качество постепенно ухудшается с ростом расхождения между используемым и актуальным наборами. Таким образом, каждый узел накапливает новые объекты, посылает накопленный пакет объектов узлу координации, а в ответ получает обновленный набор хеш-функций.

Экспериментальное исследование

Экспериментальное исследование предлагаемого подхода было произведено с использованием набора данных MovieLens 100k, выложенного в открытый доступ исследовательской лабораторией GroupLens Research [20]. Этот набор содержит 100 000 оценок, присвоенных 943 пользователями 1682 фильмам.

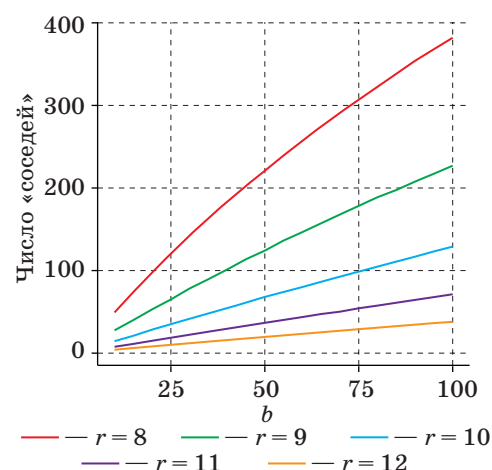
В ходе экспериментального исследования преследовались две цели. Во-первых, получить практическую информацию о количественных характеристиках подхода и оценить временную и пространственную сложность рекомендующих систем, основанных на ЛЧХ в DHT-сетях. Во-вторых, оценить качество рекомендаций по сравнению с широко распространенными альтернативными алгоритмами.

Временная и пространственная сложность

Как уже было отмечено, параметрами предлагаемого алгоритма формирования рекомендаций являются b (количество хеш-функций) и r (количество гиперплоскостей в каждой функции). Значения этих параметров оказывают существенное влияние как на время получения рекомендаций, так и на их качество.

Каждый узел помещает информацию о себе DHT b раз (по одному значению каждой из хеш-функций), следовательно, размер DHT равен Nb , где N — количество узлов, а это означает, что в среднем в узле размещено b записей DHT.

Поиск ближайших соседей требует b операций извлечения из хеш-таблицы, а значит, требует $O(b \log(N))$ взаимодействий между узлами.



■ Рис. 2. Зависимость среднего числа «соседей» от количества b хеш-функций и их размерности r

Наиболее важным параметром, хотя и не явным, а косвенным, является фактическое количество «соседей», которые обнаруживаются в результате извлечения b значений из хеш-таблиц, поскольку оно определяет количество сетевых запросов на формирование рекомендаций, и чем оно меньше (при определенном уровне качества), тем лучше.

На рис. 2 показана зависимость между параметрами b и r рекомендующей системы и средним количеством «соседей», найденных в результате поиска по хеш-таблице. Видно, что количество «соседей» увеличивается с ростом b , а темп увеличения существенно зависит от размерности хеш-функций. Это ожидаемое поведение, поскольку невысокая размерность хеш-функций и большое количество «альтернативных» хеш-функций делают процедуру поиска очень грубой. В данном исследовании предполагается, что количество хеш-функций должно быть менее 100, а количество «соседей» — менее 50.

Принимая во внимание изложенные соображения, для исследования качества рекомендаций были выбраны три конфигурации: ($r=12$, $b=100$), ($r=10$, $b=35$), ($r=8$, $b=10$), — так как в каждой из этих конфигураций количество «соседей» приблизительно равно 50.

Качество рекомендаций

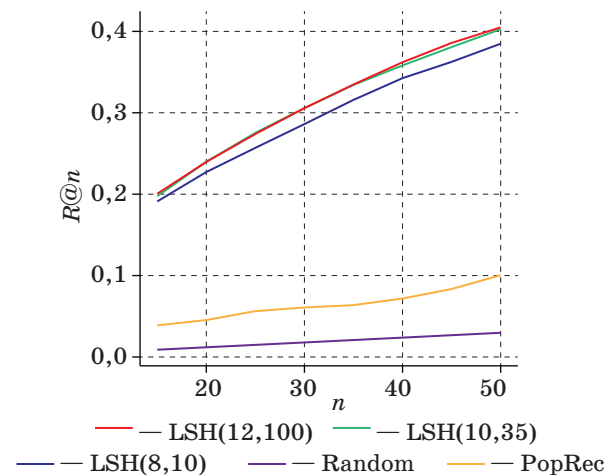
Поскольку предлагаемый алгоритм формирования рекомендаций не осуществляет предсказание самих оценок, традиционные методы оценки качества (например, среднеквадратическое отклонение между прогнозными и реальными оценками) оказываются неприменимы. Алгоритм оценивался по полноте (*recall*) — мере качества, широко распространенной при сравнении алгоритмов получения лучших n рекомендаций.

В частности, для оценки качества использовался подход, описанный в работе [21]. Набор данных с оценками был разделен на два: обучающую выборку и тестовую выборку в пропорции 80/20. Обучающая выборка использовалась для заполнения хеш-таблицы, затем для каждой высокой оценки (4 или 5) из тестовой выборки осуществлялась проверка — попадает ли соответствующий объект в список из n наиболее рекомендованных для соответствующего пользователя. Результатом такой проверки может быть либо 1 (если объект попадает в n рекомендованных), либо 0 (в противном случае). Сумма этих результатов по всему тестовому набору дает N_p . Полнота при n рекомендованных объектах (обозначаемая $R@n$) вычислялась по формуле

$$R@n = \frac{N_p}{N_H},$$

где N_H — количество высоких оценок в тестовой выборке. Другими словами, эта величина может интерпретироваться как вероятность того, что случайным образом выбранный объект с высокой оценкой действительно будет рекомендован данным алгоритмом (включен в n рекомендаций).

Полнота рекомендаций предложенной системы была сопоставлена с полнотой рекомендаций, полученных с помощью других (неперсонализированных) алгоритмов. Во-первых, случайного рекомендующего алгоритма (Random), предлагающего пользователю n случайных объектов, во-вторых, алгоритма рекомендации популярных объектов (PopRec), предлагающего объекты, получившие наибольшее количество оценок. Полнота каждого из названных рекомендующих алгоритмов при различных значениях n показана на рис. 3.



■ Рис. 3. Сравнение $R@n$ различных методов формирования рекомендаций

Все протестированные варианты формирования рекомендаций с помощью ЛЧХ дают примерно одинаковые результаты. Это можно объяснить тем, что во всех этих вариантах количество «соседей» оказывается примерно одинаковым (около 50, см. рис. 2).

Кроме того, можно увидеть, что предлагаемый алгоритм формирования рекомендаций существенно превосходит неперсонализированные рекомендующие алгоритмы по полноте.

Заключение

В статье предложена архитектура гибридной одноранговой рекомендующей системы, основанной на ЛЧХ профиля пользователей. В предлагаемой архитектуре приватность пользователей обеспечивается тем, что обмен сведениями о предпочтениях пользователей происходит только анонимизированным образом и только в виде значений хеш-функций. Было произведено экспериментальное исследование предлагаемого подхода с использованием одного из широко распространенных наборов данных и показано, что оценка полноты (*recall*) рекомендаций, формируемых с помощью предлагаемого подхода, существенно выше, чем для неперсонализированных алгоритмов формирования рекомендаций.

Однако следует отметить и некоторые ограничения предлагаемого подхода. Во-первых, из-за ограничений систем класса DHT он оказывается не подходящим для сетей с динамично изменяющейся структурой. Во-вторых, подобное устройство рекомендующей системы не подходит для областей, в которых часто появляются новые объекты (например, новостные сообщения), из-за необходимости распространения информации об объектах среди всех узлов сети.

В будущем планируется рассмотреть альтернативные решения по согласованию глобального состояния сети без нарушения одноранговой организации.

Работа выполнена при финансовой поддержке РФФИ (гранты № 13-07-00271, 13-07-00039, 14-07-00345), Президиума РАН (проект № 213) и отделения нанотехнологий и информационных технологий РАН (проект № 2.2).

Литература

1. Draid F., Pacitti E., Kemme B. P2Prec: A P2P Recommendation System for Large-Scale Data Sharing // Journal of Transactions on Large-Scale Data and Knowledge-Centered Systems (TLDKS). 2011. Vol. 3. P. 87–116.
2. Draid F., et al. P2Prec: a Social-Based P2P Recommendation System // Proc. of the 20th ACM Intern.

- Conf. on Information and Knowledge Management. 2011. P. 2593–2596.
3. Kermarrec A.-M., et al. Application of Random Walks to Decentralized Recommendation Systems // Proc. of the 14th Intern. Conf. on Principles of Distributed Systems. 2010. P. 48–63.
 4. Pitsilis G., Marshall L. A Trust-Enabled P2P Recommendation System // Proc. 15th IEEE Intern. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises. 2006. P. 59–64.
 5. Jelasity M., Montresor A., Babaoglu O. T-Man: Gossip-Based Fast Overlay Topology Construction // Computer Networks. Aug. 2009. Vol. 53. N 13. P. 2321–2339.
 6. Tveit A. Peer-to-Peer Based Recommendations for Mobile Commerce // Proc. 1st Intern. Workshop on Mobile Commerce (WMC'01), ACM. 2001. P. 26–29.
 7. Bakker A., Ogston E., van Steen M. Collaborative Filtering Using Random Neighbours in Peer-to-Peer Networks // Workshop on Complex Networks in Information & Knowledge Management. 2009. P. 67–74.
 8. Han P., et al. A Scalable P2P Recommendation System Based on Distributed Collaborative Filtering // Expert Systems with Applications. 2004. N 27(2). P. 203–210.
 9. Hecht F., et al. Radiommendation: P2P On-Line Radio with a Distributed Recommendation System // Proc. of the IEEE 12th Intern. Conf. on Peer-to-Peer Computing. 2012. P. 73–74.
 10. Rajaraman A., Ullman J. Mining of Massive Datasets. — Cambridge University Press, 2012. — 326 p.
 11. Slanley M., Casey M. Locality-Sensitive Hashing for Finding Nearest Neighbors // IEEE Signal Processing Magazine. Mar. 2008. Vol. 25. N 2. P. 128–131.
 12. Indyk P., Motwani R. Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality // STOC'98: Proc. of the 30th Symp. on Theory of Computing. 1998. P. 604–613.
 13. Charikar M. S. Similarity Estimation Techniques from Rounding Algorithms // STOC'02: Proc. of the 34th Annual ACM Symp. on Theory of Computing. 2002. P. 380–388.
 14. Korzun D., Gurtov A. Structured Peer-to-Peer Systems. Fundamentals of Hierarchical Organization, Routing, Scaling and Security. — Springer, 2013. — 366 p.
 15. Wang Q., Borisov N. Octopus: A Secure and Anonymous DHT Lookup // Proc. of the IEEE 32nd Intern. Conf. on Distributed Computing Systems. 2012. P. 325–334.
 16. Chen X., et al. SCOPE: Scalable Consistency Maintenance in Structured P2P Systems // Proc. of IEEE INFOCOM. 2005. P. 1502–1513.
 17. Hu Y., Bhuyan L. N., Feng M. Maintaining Data Consistency in Structured P2P Systems // IEEE Transactions on Parallel and Distributed Systems. 2012. Vol. 23. Iss. 11. P. 2125–2137.
 18. Oster G., et al. Data Consistency for P2P Collaborative Editing // Proc. of the 20th Anniversary Conf. on Computer Supported Cooperative Work. 2006. P. 259–268.
 19. Mastroianni C., Pirro G., Talia D. Data Consistency and Peer Synchronization in Cooperative P2P Environments. — Technical Report, 2008. — 16 p.
 20. MovieLens dataset. <http://grouplens.org/datasets/movielens/> (дата обращения: 17.05.2015).
 21. Cremonesi P., Koren Y., Turrin R. Performance of Recommender Algorithms on Top-N Recommendation Tasks // Proc. of the Fourth ACM Conf. on Recommender Systems (RecSys '10). ACM, New York, NY, USA, 2010. P. 39–46.

UDC 004.9

doi:10.15217/issn1684-8853.2015.5.91

Decentralized Recommendation System Architecture Based on Locality-Sensitive Hashing

Ponomarev A. V.^a, PhD, Senior Researcher, ponomarev@iias.spb.su^aSaint-Petersburg Institute for Informatics and Automation of RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation

Purpose: Recommendation systems are widely used in modern e-commerce systems to help users make their ways in a vast variety of offered goods and services. Most of the modern recommendation system approaches are centralized. However, centralized recommendation have two primary disadvantages: the necessity for users to share their preferences and a single point of failure. The goal of this work is developing a decentralized recommendation system which employs user similarity (collaborative filtering) but holds all the user preferences only on the user's network node. **Results:** An architecture is proposed for a decentralized recommendation system. It includes a structured peer-to-peer network in which each node corresponds to one user and stores this user's preferences, and a special node used for the coordination of peer-to-peer nodes in some scenarios. To find users with similar interests and, in the same time, restrict the sharing of preferences, a locality-sensitive hashing is employed. For a higher level of privacy, the network uses an anonymization scheme. **Practical relevance:** The proposed approach is universal, as it relies only on ratings, and can be used to build collaborative filtering systems in various domains.

Keywords — Locality-Sensitive Hashing, Peer-To-Peer, Recommendation Systems, Collaborative Filtering.

References

1. Draidi F., Pacitti E., and Kemme B. P2Prec: A P2P Recommendation System for Large-Scale Data Sharing. *Journal of*

Transactions on Large-Scale Data and Knowledge-Centered Systems (TLDKS), 2011, vol. 3, pp. 87–116.

2. Draidi F., et al. P2Prec: A Social-Based P2P Recommendation System. *Proc. of the 20th ACM Intern. Conf. on Information and Knowledge Management*, 2011, pp. 2593–2596.
 3. Kermarrec A.-M., et al. Application of Random Walks to Decentralized Recommendation Systems. *Proc. of the 14th Intern. Conf. on Principles of Distributed Systems*, 2010, pp. 48–63.
 4. Pitsilis G., Marshall L. A Trust-Enabled P2P Recommendation System. *Proc. 15th IEEE Intern. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises*, 2006, pp. 59–64.
 5. Jelasity M., Montresor A., Babaoglu O. T-Man: Gossip-Based Fast Overlay Topology Construction. *Computer Networks*, August 2009, vol. 53, no. 13, pp. 2321–2339.
 6. Tveit A. Peer-to-Peer Based Recommendations for Mobile Commerce. *Proc. 1st Intern. Workshop on Mobile Commerce (WMC'01)*, ACM, 2001, pp. 26–29.
 7. Bakker A., Ogston E., and van Steen M. Collaborative Filtering Using Random Neighbours in Peer-to-Peer Networks. *Workshop on Complex Networks in Information & Knowledge Management*, 2009, pp. 67–74.
 8. Han P., et al. A Scalable P2P Recommendation System Based On Distributed Collaborative Filtering. *Expert Systems with Applications*, 2004, no. 27(2), pp. 203–210.
 9. Hecht F., et al. Radiommendation: P2P On-Line Radio with a Distributed Recommendation System. *Proc. of the IEEE 12th Intern. Conf. on Peer-to-Peer Computing*, 2012, pp. 73–74.
 10. Rajaraman A., Ullman J. *Mining of Massive Datasets*. Cambridge University Press, 2012. 326 p.
 11. Stanley M., Casey M. Locality-Sensitive Hashing for Finding Nearest Neighbors. *IEEE Signal Processing Magazine*, March 2008, vol. 25, no. 2, pp. 128–131.
 12. Indyk P., Motwani R. Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality. *STOC'98 Proc. of the 30th Symp. on Theory of Computing*, 1998, pp. 604–613.
 13. Charikar M. S. Similarity Estimation Techniques from Rounding Algorithms. *STOC'02 Proc. of the 34th annual ACM Symp. on Theory of Computing*, 2002, pp. 380–388.
 14. Korzun D., Gurtov A. *Structured Peer-to-Peer Systems. Fundamentals of Hierarchical Organization, Routing, Scaling and Security*. Springer, 2013. 366 p.
 15. Wang Q., Borisov N. Octopus: A Secure and Anonymous DHT Lookup. *Proc. of the IEEE 32nd Intern. Conf. on Distributed Computing Systems*, 2012, pp. 325–334.
 16. Chen X., et al. SCOPE: Scalable Consistency Maintenance in Structured P2P Systems. *Proc. of IEEE INFOCOM*, 2005, pp. 1502–1513.
 17. Hu Y., Bhuyan L. N., and Feng M. Maintaining Data Consistency in Structured P2P Systems. *IEEE Transactions on Parallel and Distributed Systems*, 2012, vol. 23, iss. 11, pp. 2125–2137.
 18. Oster G., et al. Data Consistency for P2P Collaborative Editing. *Proc. of the 20th Anniversary Conf. on Computer Supported Cooperative Work*, 2006, pp. 259–268.
 19. Mastroianni C., Pirro G., and Talia D. *Data Consistency and Peer Synchronization in Cooperative P2P Environments*. Technical Report, 2008. 19 p.
 20. MovieLens dataset. Available at: <http://grouplens.org/datasets/movielens/> (accessed 17 May 2015).
 21. Cremonesi P., Koren Y., Turrin R. Performance of Recommender Algorithms on Top-N Recommendation Tasks. *Proc. of the fourth ACM Conf. on Recommender Systems (RecSys '10)*, ACM, New York, NY, USA, 2010, pp. 39–46.
-

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию. Рукописи не возвращаются.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

ЭВРИСТИЧЕСКИЙ АЛГОРИТМ РАСЧЕТА РАЗМЕРОВ ПАМЯТИ В МНОГОУРОВНЕВОЙ СИСТЕМЕ ХРАНЕНИЯ

Д. А. Маличенко^{а, 1}, ведущий программист

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Введение: рассматривается многоуровневая блочная система хранения данных. Уровни отличаются типом памяти и скоростью работы с данными. Для каждого приложения, работающего с системой, задано требуемое качество обслуживания, выражаемое в виде времени отклика системы на запрос. Распределение адресного пространства приложений между уровнями памяти влияет на время отклика системы. Целью работы является автоматическое распределение областей памяти приложений в зависимости от входного потока заявок таким образом, чтобы требования на качество обслуживания были выполнены. **Результаты:** сформулирована оптимизационная задача расчета размеров памяти приложений на всех уровнях системы и предложен эвристический алгоритм ее решения. Используются такие характеристики потока запросов, как среднее значение доли запросов, попадающих в кэш, и распределение частот запросов к блокам данных на дисках. Оценка работы алгоритма вычислена с помощью имитационного моделирования системы хранения, на которую подавался поток заявок, снятый ранее с реальных систем хранения. Результаты моделирования демонстрируют, что предложенный алгоритм в некоторых случаях позволяет повысить эффективность работы системы хранения до 10 %. **Практическая значимость:** предложенный алгоритм может улучшить качество обслуживания в многоуровневой системе хранения данных.

Ключевые слова — многоуровневая система хранения, управление многоуровневой системой хранения, Storage Area Network, качество обслуживания, моделирование.

Введение

Системы хранения данных очень разнообразны по уровню сложности. Для некоторых задач обычные жесткие диски, которые используются в персональных компьютерах, не подходят. Например, у сервисов по бронированию и покупке билетов или у социальных сетей возникают дополнительные требования к системам хранения данных. Требуется хранить больше данных, одновременно обрабатывать заявки от большого числа клиентов, а также быстро работать с данными. Для таких целей используются специализированные сложные системы хранения, которые могут представлять собой целую сеть. В работе рассматривается пример такой системы хранения данных. Для увеличения скорости работы системы используются диски разного типа. Применение только очень быстрых дисков значительно увеличивает стоимость всей системы. Обычно комбинируют диски с разной скоростью работы. В этом случае более востребованные данные хранятся на более быстрых дисках, менее востребованные данные хранятся на медленных дисках. Со временем востребованность данных меняется, и управляющая система должна перекладывать данные между дисками разного типа.

¹ Научный руководитель — профессор, проректор по науке, доктор технических наук, директор Института информационных систем и защиты информации, ведущий кафедрой безопасности информационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения Е. А. Крук.

С системой хранения работает множество клиентов — приложений. У каждого приложения свой объем востребованных данных, следовательно, пропорции, в которых данные распределены между дисками разного типа, у каждого приложения свои. Одному приложению может требоваться больше быстрой памяти для обеспечения заданной скорости работы, другому меньше, а со временем эта ситуация может поменяться в обратную сторону. Если объем быстрой памяти, выделенный приложению, слишком мал, то все наиболее востребованные данные приложения не поместятся в быструю память, и часть данных будет считываться с медленных дисков, что приведет к уменьшению средней скорости работы с данными для приложения. Кроме дискового пространства у каждого приложения есть своя область в кэш-памяти. Требуемые размеры этих областей также разные для разных приложений и зависят от характеристик потока запрашиваемых данных. В данной работе рассматривается задача выбора для каждого приложения размера выделяемой памяти в кэше и на дисках всех типов.

Многоуровневые системы хранения известны давно. В работе [1] дано функциональное описание такой системы, а в работе [2] предложен алгоритм перемещения блоков данных между дисками разных типов. В работе [3] для идеализированной модели системы хранения представлены аналитические выражения для расчета требуемого размера кэша, а также для размеров дисков и оптимального количества уровней системы хранения. Оптимизация размеров кэш-памяти рас-

смотрена в работах [4–6], а распределение места на дисках — в работах [7, 8]. В работе [7] система состоит из двух уровней памяти: SSD-дисков и обычных дисков, причем SSD-память работает в режиме кэша.

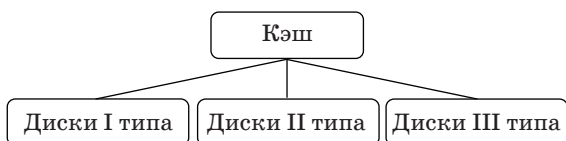
В настоящей работе, в отличие от многих других, рассматривается изоляция приложений друг от друга в кэше и на дисках. Решения алгоритма распространяются не только на диски, но и на кэш, для которого вычисляются требуемые хит-рейты.

Многоуровневая система хранения данных

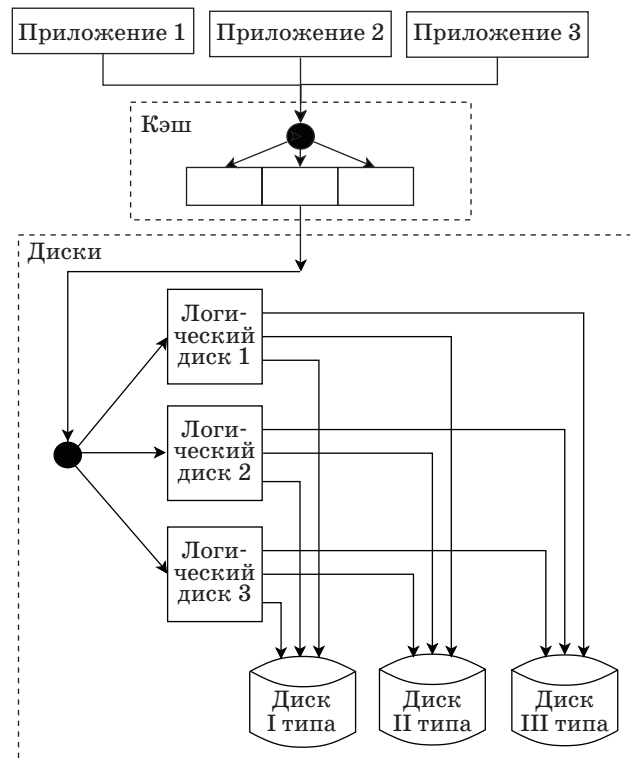
Опишем систему хранения с четырьмя типами памяти. Имеется кэш-память и три типа дисков (рис. 1), которые отличаются объемом и скоростью работы. Запрашиваемый блок данных сначала проверяется на наличие в кэше. Если блок в кэше отсутствует, то он считывается с того диска, на котором записан.

Рассмотрим сценарий, в котором у каждого приложения, работающего с системой, есть своя независимая область в кэше и свой логический диск, при этом физически пространство логического диска распределено между дисками всех типов. Более подробно работа системы проиллюстрирована на рис. 2.

Когда приложение 1 запрашивает блок данных, то система хранения сначала ищет его в выделенной для приложения 1 области кэша. Если блок найден, то время чтения блока равно времени чтения из кэша. Если блок отсутствует в кэше, то он считывается с диска. Приложение работает с логическим диском. На более низком уровне системы хранения лежат физические диски разного типа. Блок данных может лежать либо на быстрых дисках, либо на медленных. Примером быстрых дисков могут служить диски на flash-памяти (SSD-диски), примером медленных дисков могут служить SATA-диски. Расположение блоков данных постоянно меняется системой. Будем считать, что раз в час система перемещает блоки данных в соответствии с вероятностями их запросов. Все блоки сортируются по популярности, и самые популярные записываются на самые быстрые диски, пока место, выделенное для данного приложения, не закончится; далее блоки записы-



■ Рис. 1. Упрощенная схема многоуровневой системы хранения



■ Рис. 2. Схема работы многоуровневой системы хранения с разделением приложений

ваются на более медленные диски. Статистика по вероятностям запросов собирается в течение часа до следующего перемещения блоков.

Определение основных зависимостей и постановка задачи

Пусть время доступа к блоку данных в кэше равно t_0 , а времена доступа к дискам t_1, t_2, t_3 и $t_0 \ll t_1 \ll t_2 \ll t_3$. Пусть объемы памяти в кэше и на дисках соответственно равны S_0, S_1, S_2, S_3 . С системой работает n приложений $A_1, \dots, A_n$. Пусть объем памяти, выделенный для i -го приложения на j -м уровне памяти, равен s_j^i . Под уровнем понимается тип памяти. Тогда

$$\sum_{i=1}^n s_j^i \leq S_j. \quad (1)$$

Пусть p_j^i — вероятность того, что блок данных, запрошенный i -м приложением, окажется на j -м уровне памяти. Тогда среднее время доступа к блоку данных для i -го приложения

$$\bar{T}_i = p_0^i t_0 + p_1^i t_1 + p_2^i t_2 + p_3^i t_3, \quad (2)$$

где $p_0^i + p_1^i + p_2^i + p_3^i = 1$.

Пусть заданы ограничения на средние времена доступа к блоку данных для каждого приложения:

$$\bar{T}_i \leq R_i. \quad (3)$$

Следовательно, присутствует ограничение на вероятности p_j^i

$$p_0^i t_0 + p_1^i t_1 + p_2^i t_2 + p_3^i t_3 \leq R_i. \quad (4)$$

Вероятности p_j^i связаны с объемами памяти s_j^i . Чем больше памяти приложение использует на j -м уровне относительно других уровней памяти, тем выше вероятность p_j^i . Обозначим эту зависимость так:

$$p_0^i = f_c^i(s_0^i);$$

$$p_j^i = f_d^i(s_j^i, p_0^i), \quad 1 \leq j \leq 3,$$

где $f_c^i()$ и $f_d^i()$ — функции зависимости между размером занимаемой памяти и вероятностью попадания на нее для кэш-памяти и дисков соответственно.

Вероятность p_0^i называют хит-рейтом. Важной характеристикой кэш-памяти является стековое расстояние [9] — это количество уникальных блоков данных, запрошенных между запросом одного и того же блока, плюс один. Например, если блоки запрашивались в последовательности А, В, С, D, Е, D, С, В, D, А, то стековое расстояние для второго появления блока А равно пяти. Функция распределения стековых расстояний $F_{sd}^i(x)$ дает вероятность того, что стековое расстояние для следующего запрошенного блока будет меньше либо равно x . Это означает, что если иметь в распоряжении кэш размером x блоков, то хит-рейт кэша будет равен $F_{sd}^i(x)$. Для i -го приложения этот факт можно записать таким образом:

$$p_0^i = F_{sd}^i(s_0^i).$$

Заметим, что $f_c^i(s_0^i) = F_{sd}^i(s_0^i)$.

Рассмотрим теперь $F_{pop}^i(x)$ — функцию распределения вероятностей того, что следующий запрошенный i -м приложением блок данных будет находиться среди первых x блоков с наивысшей вероятностью появления. Иногда $F_{pop}^i(x)$ называют популярностью адресов [9]. Чтобы получить эту функцию для потока запрашиваемых блоков данных, необходимо посчитать частотные вероятности каждого блока. После этого отсортировать пары значений (номер блока, вероятность) по убыванию вероятности. Перенумеруем блоки по порядку после сортировки, начиная от самого вероятного. Функция, задаваемая этими парами, где аргумент — номер блока по порядку, а значением является вероятность, будет функцией плотности вероятности блока, а построенная по ней функция распределения будет искомой $F_{pop}^i(x)$. Если имеется три диска разных типов с размерами S_1, S_2, S_3 , и данные упорядочены по степени востребованности так, что самые востребованные хранятся на диске 1, а самые невостребованные — на диске 3, то вероятность

того, что данные будут запрошены с диска 1, равна

$$p_1^i = \frac{F_{pop}^i(s_1^i)}{1 - p_0^i} = f_d^1, \quad (5)$$

вероятность запроса данных с диска 2

$$p_2^i = \frac{F_{pop}^i(s_2^i + s_1^i) - F_{pop}^i(s_1^i)}{1 - p_0^i} = f_d^2, \quad (6)$$

а вероятность запроса данных с диска 3

$$p_3^i = \frac{F_{pop}^i(s_3^i + s_2^i + s_1^i) - F_{pop}^i(s_2^i + s_1^i)}{1 - p_0^i} = f_d^3. \quad (7)$$

Таким образом, функция f_d^i для дисков разных уровней будет отличаться. В рассматриваемой системе хранения приложению выделена только часть s_j^i каждого диска, т. е. $s_j^i < S_j$. Перепишем уравнения (5)–(7) другим образом:

$$s_1^i = F_{pop}^{i-1}(p_1^i(1 - p_0^i)); \quad (8)$$

$$s_2^i + s_1^i = F_{pop}^{i-1}((p_2^i + p_1^i)(1 - p_0^i)); \quad (9)$$

$$s_3^i + s_2^i + s_1^i = F_{pop}^{i-1}((p_3^i + p_2^i + p_1^i)(1 - p_0^i)). \quad (10)$$

При этом на вероятности p_j^i имеется ограничение (4), а на s_j^i — ограничение (1).

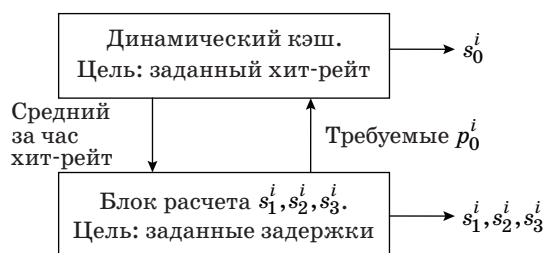
Все необходимые зависимости определены, перейдем к постановке задачи. Пусть известны параметры $S_0, S_1, S_2, S_3, t_0, t_1, t_2, t_3, n, R_i = 1, \dots, n$, требуется найти такие s_j^i , чтобы выполнялись заданные ограничения $R_i = 1, \dots, n$.

Кроме заданных параметров, в выражениях (8)–(10) присутствуют функции $F_{sd}^i(x)$ и $F_{pop}^i(x)$, которые меняются в зависимости от потока запросов, и для реальной системы их можно получить только численно, они задаются таблицей. Это приводит к тому, что известные методы решения задач оптимизации в данном случае не применимы. По этой причине нами предлагается эвристический алгоритм решения задачи.

Описание предлагаемого решения

Функционально предлагается заменить исходный кэш со статичным разделением областей между приложениями на описанный в работе [6] динамический кэш, который меняет границы областей приложений в зависимости от заданного для каждого приложения хит-рейта. В настоящей работе вводится дополнительный блок (рис. 3), который раз в час будет пересчитывать для каждого приложения размеры областей, выделяемых системой на дисках разного типа.

Значения s_0^i вычисляются динамическим кэшем [6], чтобы достичь хит-рейтов, заданных для



■ Рис. 3. Функциональная схема предлагаемого решения

приложений. В предлагаемой схеме хит-рейты задает блок расчета s_1^i, s_2^i, s_3^i . В начале работы системы размеры s_0^i задаются некоторым типовым образом, например, одинаковыми или пропорционально требованиям на отклик системы.

Рассмотрим алгоритм расчета s_1^i, s_2^i, s_3^i . Как упомянуто выше, алгоритм вычисляет границы областей один раз в час. За час собирается статистика, вычисляется $F_{pop}^i(x)$, и вместе с текущими значениями границ по выражениям (5)–(7) могут быть рассчитаны вероятности p_1^i, p_2^i, p_3^i . Используя эти вероятности и p_0^i , которую можно посчитать в кэше, с помощью выражения (2) вычисляем $\bar{T}_i$. Далее все $\bar{T}_i$ разделим на два списка. В первый список попадут $\bar{T}_i \geq R_i$, а во второй список — $\bar{T}_i < \delta R_i$, где δ характеризует необходимый запас по требуемой задержке и представляет собой долю от R_i . В первый список $T1$ попадут задержки тех приложений, для которых требование R_i не выполнилось, а во второй список $T2$ — задержки тех, для которых они были выполнены с избытком. Если эти списки не пустые, то далее задача алгоритма состоит в таком изменении s_1^i, s_2^i, s_3^i , при котором для всех приложений требования R_i станут удовлетворяться. Для этого сначала списки $T1$ и $T2$ сортируются по возрастанию. Далее для приложений, чьи $\bar{T}_i$ оказались в начале списков, происходит изменение размеров областей на дисках. Изменения производятся пошагово, по одному слоту, начиная с самой быстрой памяти. Сначала s_1^i приложения с избытком уменьшается на один слот, а s_1^i приложения с недостатком увеличивается на один слот. После этого по формуле (2) снова пересчитываются $\bar{T}_i$ для двух приложений. В таблице представлены различные варианты значений $\bar{T}_i$ и дальнейшие шаги.

Параметр α необходим для того, чтобы приложение с запасом по $\bar{T}_i$ не стало приложением, не удовлетворяющим требованию. Если выполнилось условие 2, то дальше выбирается следующая пара приложений с недостатком и избытком $\bar{T}_i$, причем приложение с избытком может остаться тем же, что было на предыдущем шаге. Если выполнилось условие 3, то резерв быстрой памяти у приложения с избытком закончился. В этом случае после шага назад приложение

■ Возможные события в результате перераспределения памяти

№	Условие	Действие
1	$\bar{T}_x > R_x, \bar{T}_y < \alpha R_y$	Продолжать увеличивать s_1^x и уменьшать s_1^y
2	$\bar{T}_x < R_x, \bar{T}_y < \alpha R_y$	Остановить перераспределение s_1^x
3	$\bar{T}_y > \alpha R_y$	Сделать шаг назад, т. е. уменьшить s_1^x и увеличить s_1^y

с недостатком все еще не удовлетворяет требованию. Далее необходимо повторить процедуру перераспределения для следующего уровня памяти s_2^i . Если и в нем будет такой же результат, то выбирается следующее приложение с избытком, и процедура перераспределения памяти повторяется с уровня s_1^i . Далее либо закончатся все приложения с недостатком быстрой памяти, что означает успешное завершение алгоритма, либо закончатся приложения с избытком, а с недостатком останутся. В последнем случае система не сможет удовлетворить текущим требованиям на задержку.

Работа алгоритма основана на предположении, что во время функционирования системы присутствуют достаточно длинные периоды стабильности $F_{pop}^i(x)$, тогда, собрав за определенное время статистику, можно рассчитывать на то, что новые значения s_j^i будут актуальны.

Для того чтобы быстрее реагировать на изменения потока запросов, в алгоритме предусмотрена процедура пересчета требуемого хит-рейта. Она работает каждые 10 мин, в то время, когда пересчет s_1^i, s_2^i, s_3^i не производится. По текущим значениям s_1^i, s_2^i, s_3^i с помощью формул (5)–(7) вычисляются p_1^i, p_2^i, p_3^i . Далее из этих формул можно выразить вероятность p_0^i , которая и является требуемым хит-рейтом. Найденное значение передается в блок управления кэшем, который должен в результате работы выдерживать заданные p_0^i .

Результаты моделирования

Определим коэффициент Q , по которому предложенный алгоритм будет оцениваться и сравниваться с системой без пересчета размеров областей на дисках. Для этого введем вспомогательный коэффициент

$$q_i = \frac{N\{\bar{T}_i \leq R_i\}}{N}, \quad (11)$$

где $N\{\bar{T}_i \leq R_i\}$ — количество десятиминутных интервалов, в течение которых среднее время считывания блока данных не превышало заданное для приложения i , а N — общее количество

десятиминутных интервалов работы системы. Тогда

$$Q = \frac{1}{n} \sum_{i=1}^n q_i, \quad (12)$$

откуда видно, что Q характеризует работу системы в целом по всем приложениям.

Для вычисления коэффициента Q воспользуемся имитационным моделированием. В качестве модели кэша будем использовать LRU (*Least Recently Used*) стек [9]. В случае с постоянными размерами областей памяти в кэше будем использовать несколько LRU-стеков фиксированного размера. В случае с изменяемыми размерами областей размер будут пересчитываться [6]. В качестве модели дисков будем использовать таблицу, в которой записано, к диску какого типа относится каждый блок данных. Каждый час блоки будут перемещаться между дисками разного типа согласно $F_{pop}^i(x)$, как описано выше. В качестве потока заявок использовались логи реальных систем хранения [10]. На рис. 4 представлены пары коэффициентов Q для одного синтезированного лога и для нескольких общедоступных логов реальных систем хранения. Первый коэффициент получен для системы с размерами s_j^i , выставленными пропорционально требованиям на задержку (статические границы). Второй коэффициент получен при работе системы с предложенным алгоритмом пересчета s_j^i (динамические границы).

Логи Financial1, Financial2 сняты с систем обработки транзакций в реальном времени в финансовых организациях. Логи WebSearch2, WebSearch3 собраны с известных поисковых систем [10]. Из рисунка видно, что наибольший выигрыш получается на синтезированном логе. Как было отмечено, предлагаемый алгоритм опирается на предположение о наличии достаточно длинных периодов времени, на которых функция $F_{pop}^i(x)$ постоянна. В синтезированном логе это условие строго выполняется, поэтому выигрыш наибольший. В реальности поведение $F_{pop}^i(x)$ непредсказуемо. Кроме предсказуемости функции популярности для успеха алгоритма необходимо, чтобы приложениям требовались разные объемы памяти для выполнения заданных ограничений и чтобы эти объемы достаточно отличались от выделенного изначально объема. На логах

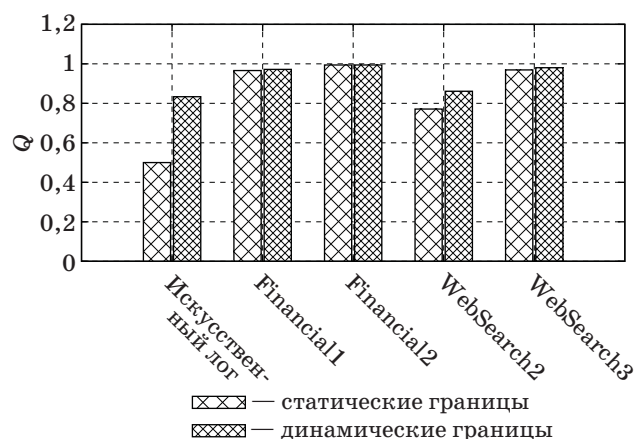


Рис. 4. Результаты моделирования предложенного алгоритма

Financial1 и WebSearch3 наблюдается незначительный выигрыш предложенного алгоритма. На логе Financial2 разницы нет. Это связано с тем, что свободных для перераспределения ресурсов не было и выдерживались изначально выставленные границы. На логе WebSearch2 выигрыш предложенного алгоритма составляет 12 %. Можно также отметить, что благодаря введенным защитным параметрам α и δ , на всех рассмотренных логах критерий Q не стал хуже по сравнению с исходной системой (со статическими границами). Также были исследованы логи и других систем. На них выигрыш составил 1–10 %.

Заключение

Работа предложенного в статье алгоритма перераспределения места на дисках основана на предположении о стабильности распределения популярностей адресов. Если это выполняется, то алгоритм дает ощутимое преимущество, что показано с помощью моделирования на искусственном потоке запросов адресов. При использовании потока с реальными системами данное предположение может не выполняться, однако моделирование показало, что довольно часто и на реальных потоках запросов алгоритм дает выигрыш до 10 %. Это позволяет сделать вывод, что алгоритм можно использовать как помощника при конфигурировании системы хранения.

Литература

1. Morenoff E., McLean J. Application of Level Changing to a Multilevel Storage Organization // Communications of the ACM. Mar. 1967. Vol. 10. Iss. 3. P. 149–154. doi:10.1145/363162.363183
2. Ver Hoef E. W. Design of a Multi-Level File Management System // Proc. of the 1966 21st National Conf. 1966. P. 75–86. doi: 10.1145/800256.810684
3. Chow C. K. Determination of Cache's Capacity and its Matching Storage Hierarchy // IEEE Transactions on Computers. Feb. 1976. Vol. C-25. Iss. 2. P. 157–164. doi: 10.1109/TC.1976.5009230
4. Lai P., Fan R. Makespan-Optimal Cache Partitioning // IEEE 21st Intern. Symp. on Modeling, Analysis

- & Simulation of Computer and Telecommunication Systems (MASCOTS), 14–16 Aug. 2013. P. 202–211. doi: 10.1109/MASCOTS.2013.28
5. Yuan L. A Locality-based Performance Model for Load-and-Compute Style Computation // IEEE Intern. Conf. on Cluster Computing (CLUSTER), 24–28 Sept. 2012. P. 566–571. doi: 10.1109/CLUSTER.2012.25
 6. Дужин В. С. Алгоритм перераспределения кэша между приложениями в системе хранения данных // Научная сессия ГУАП: сб. докл.: в 3 ч. Ч. I. Технические науки. СПб.: ГУАП, 2012. С. 73–75.
 7. Liu X., Salem K. Hybrid Storage Management for Database Systems // Proc. of the VLDB Endowment, June 2013. Vol. 6. Iss. 8. P. 541–552. doi: 10.14778/2536354.2536355
 8. Canim M., et al. An Object Placement Advisor for DB2 Using Solid State Storage // Proc. of the VLDB Endowment. Aug. 2009. Vol. 2. Iss. 2. P. 1318–1329.
 9. Feitelson D. G. Workload Modeling for Computer Systems Performance Evaluation. — NY.: Cambridge University Press, 2015. — 564 p.
 10. UMass Trace Repository. <http://traces.cs.umass.edu/index.php/Storage/Storage> (дата обращения: 10.09.2015).

UDC 681.324

doi:10.15217/issn1684-8853.2015.5.100

Heuristic Algorithm for Calculating the Size of Tiers in a Hierarchical Storage SystemMalichenko D. A.^a, Senior Programmer, dml@vu.spb.ru^aSaint-Petersburg State University of Aerospace Instrumentations, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: A hierarchical storage system which serves many applications is discussed. Levels in the hierarchy differ in their memory type and response time. For each application, a certain quality of the service is assigned in terms of latency. The allocation of the address space among applications and memory levels affects the response time of the system. The purpose of the work is automatic memory allocation at every level taking into account the request flow characteristics in order to meet the service quality demands. **Results:** A memory partitioning optimization problem is formulated, and a heuristic algorithm is proposed for its solution. The cache hit rate and distribution function of popularity of the requested memory blocks are used as flow characteristics. An assessment of the storage system efficiency is calculated by simulating a storage system fed by a request flow saved from real storage systems. The calculated assessments show up to 10% gain for the proposed algorithm. **Practical relevance:** The proposed algorithm can improve the service quality in a hierarchical storage system.

Keywords — Hierarchical Storage System, Hierarchical Storage Management, Storage Area Network, Quality of Service, Simulation.

References

1. Morenoff E., McLean J. Application of Level Changing to a Multilevel Storage Organization. *Communications of the ACM*, March 1967, vol. 10, iss. 3, pp. 149–154. doi:10.1145.363162.363183
2. Ver Hoef E. W. Design of a Multi-Level File Management System. *Proc. of the 1966 21st National Conf.*, 1966, pp. 75–86. doi: 10.1145/800256.810684
3. Chow C. K. Determination of Cache's Capacity and its Matching Storage Hierarchy. *IEEE Transactions on Computers*, February 1976, vol. C-25, iss. 2, pp. 157–164. doi: 10.1109/TC.1976.5009230
4. Lai P., Fan R. Makespan-Optimal Cache Partitioning. *IEEE 21st Intern. Symp. on Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS)*, 14–16 August 2013, pp. 202–211. doi: 10.1109/MASCOTS.2013.28
5. Yuan L. A Locality-based Performance Model for Load-and-Compute Style Computation. *IEEE Intern. Conf. on Cluster Computing (CLUSTER)*, 24–28 September 2012, pp. 566–571. doi: 10.1109/CLUSTER.2012.25
6. Duzhin V. S. Cache Partitioning Algorithm in a Storage System. *Nauchnaya sessiya GUAP* [Proceedings of Scientific Session of SUAI]. Saint-Petersburg, GUAP Publ., 2012, vol. III "Technical science", pp. 73–75 (In Russian).
7. Liu X., Salem K. Hybrid Storage Management for Database Systems. *Proc. of the VLDB Endowment*, June 2013, vol. 6, iss. 8, pp. 541–552. doi: 10.14778/2536354.2536355
8. Canim M., Mihaila G. A., Bhattacharjee B., Ross K. A., Lang C. A. An Object Placement Advisor for DB2 Using Solid State Storage. *Proc. of the VLDB Endowment*, August 2009, vol. 2, iss. 2, pp. 1318–1329.
9. Feitelson D. G. *Workload Modeling for Computer Systems Performance Evaluation*. New York, Cambridge University Press, 2015. 564 p.
10. UMass Trace Repository. Available at: <http://traces.cs.umass.edu/index.php/Storage/Storage> (accessed 10 September 2015).

ИССЛЕДОВАНИЕ НЕЛИНЕЙНЫХ ИСКАЖЕНИЙ РАДИОСИГНАЛА В СИСТЕМЕ CDMA С КВАДРАТУРНО-АМПЛИТУДНОЙ МОДУЛЯЦИЕЙ

В. Э. Гуревич^а, канд. техн. наук, профессор

С. Г. Егоров^а, соискатель, ассистент

^аСанкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича, Санкт-Петербург, РФ

Постановка проблемы: нелинейные искажения группового сигнала в цифровых радиосистемах абонентского доступа с прямым расширением спектра, кодовым разделением каналов и квадратурно-амплитудной модуляцией приводят к увеличению вероятности цифровых ошибок и ухудшению качества индивидуальных абонентских каналов. **Цель:** установление количественной зависимости вероятности цифровых ошибок в абонентских каналах от степени нелинейных искажений групповых видео- и радиосигналов. Знание этой зависимости необходимо для обоснования требований к амплитудным характеристикам усилителей, входящих в состав многоканальных видео- и радиотрактов. **Методы:** реальный групповой сигнал представлен в виде суммы трех компонентов — неискаженной части сигнала, коррелированной с ним мультипликативной и некоррелированной аддитивной помехи. Построена и применена компьютерная модель передаваемого сигнала, учитывающая, в отличие от традиционных методов анализа, не только статистические, но и комбинаторные свойства сложного группового сигнала с кодовым разделением. Приемлемая погрешность оценки помехоустойчивости достигается путем аппроксимации реального распределения смеси сигнала с помехами при помощи полиномов Эрмита и рядов Грама — Шарлье. **Результаты:** показано, что обычное гауссово представление помех и линеаризация характеристик нелинейных элементов приводят к завышенным оценкам помехоустойчивости. Анализ комбинаторики группового сигнала позволил оценить как раздельное, так и совокупное влияние каждого из параметров нелинейных трактов на реальную помехоустойчивость системы. Законы распределения смеси канальных функций Радемахера — Уолша с шумом в нелинейном тракте, а также конкретные результаты расчета вероятности ошибок представлены в аналитической и графической форме. **Практическая значимость:** результаты исследования и вытекающие из них рекомендации по заданию технических требований к групповым трактам систем с кодовым разделением позволяют повысить помехоустойчивость абонентских каналов доступа и могут быть использованы в процессах теоретических исследований, технической разработки, производства и внедрения новых высокоскоростных цифровых систем передачи.

Ключевые слова — система радиодоступа DS-CDMA, ортогональные функции Радемахера — Уолша, квадратурно-амплитудная модуляция, нелинейные усилители радиосигнала и видеосигнала, комбинаторные свойства группового сигнала, коррелированные и некоррелированные помехи, отношение сигнал/шум, вероятность битовой ошибки.

Введение

Одна из важных проблем разработки и проектирования радиосистем множественного доступа с кодовым разделением каналов, прямым расширением спектра (*Direct Sequence Code Division Multiple Access* — DS-CDMA) и квадратурно-амплитудной модуляцией (QAM) состоит в задании технических требований к линейности групповых трактов. Завышенные требования приводят к значительному усложнению системы, заниженные — к ухудшению качества связи.

Исследованию помехоустойчивости группового видеотракта системы CDMA посвящены работы [1, 2]. Однако нелинейные искажения в радиотракте также могут приводить к возникновению цифровых ошибок, поэтому важно исследовать обе причины ухудшения реальной помехоустойчивости в совокупности.

В настоящей работе помехоустойчивость системы DS-CDMA с QAM анализируется комплексно, с учетом искажений, возникающих как в видео-, так и в радиотракте. На основе проведенного исследования сформулированы требования к от-

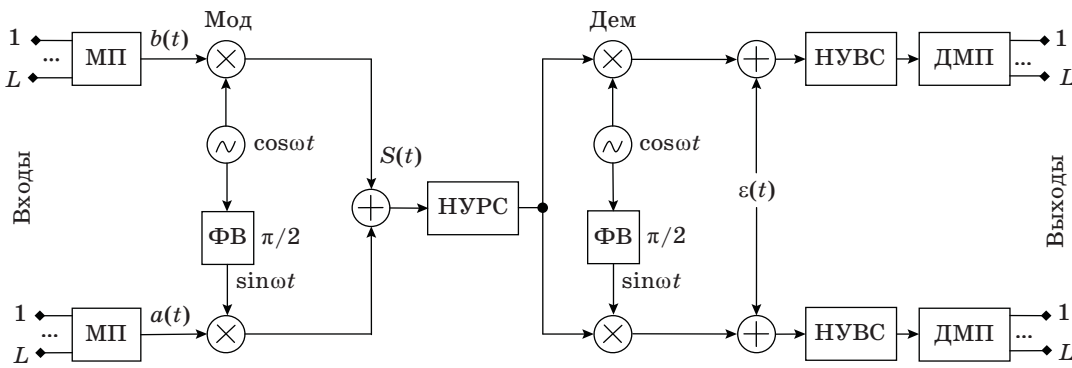
ношению сигнал/аддитивный шум и амплитудным характеристикам (АХ) усилителей многоканального (группового) сигнала.

Модель группового радиосигнала

В системах DS-CDMA с QAM цифровой сигнал каждого абонента модулирует присвоенную ему двоичную кодовую последовательность из ортогонального базиса размерности N , например функцию Радемахера — Уолша. Каждый интервал ортогональности (ИО) состоит из N элементов — чипов. В отдельном взятом ИО случайные значения чипов группового видеосигнала на входах модуляторов независимых друг от друга синфазного I и квадратурного Q компонентных стволов (рис. 1) можно записать как

$$a(t) = \sum_{m=0}^L D_I^{(m)}(t) w^{(m)}(t);$$

$$b(t) = \sum_{n=0}^L D_Q^{(n)}(t) w^{(n)}(t),$$



■ **Рис. 1.** Структурная схема системы DS-SS-SSM с QAM: МП — мультиплексоры; Мод — модуляторы; Дем — демодуляторы; ФВ — фазовращатели; НУРС и НУВС — нелинейные усилители радио- и видеосигнала соответственно; ДМП — демультиплексоры

где $L \leq N$ — количество каналов, задействованных в стволе¹; $D_1^{(m)}$ и $D_Q^{(n)}$ — информационные символы (условно говоря, «биты») $+1$ или -1 , поступающие от m -го и n -го абонентов соответствующего ствола; $w^{(m)}(t)$ и $w^{(n)}(t)$ — функции Радемахера — Уолша, закрепленные за этими абонентами и тоже принимающие значения $+1$ или -1 в каждом чипе.

Случайные параметры группового радиосигнала $S(t) = A(t)\cos[\omega t + \varphi(t)]$ в пределах ИО: амплитуда $A(t) = \sqrt{a^2(t) + b^2(t)}$ и начальная фаза $\varphi(t) = \arctg[b(t)/a(t)]$, — постоянные в пределах каждого чипа, но зависящие от номера $r = 1, 2, \dots, N$ чипа в ИО, при $L \geq 2$ практически не коррелированы между собой.

Распределение по множеству величин чипов (рис. 2), принимающих в стволах **I** и **Q** нормированные дискретные значения $a_i = 2i - L$ и $b_k = 2k - L$ соответственно, где $i, k = 0, 1, \dots, L$, подчиняется биномиальному закону:

$$P_a(a_i) = C_L^2 q^{\frac{L+a_i}{2}} (1-q)^{\frac{L-a_i}{2}}; \quad (1a)$$

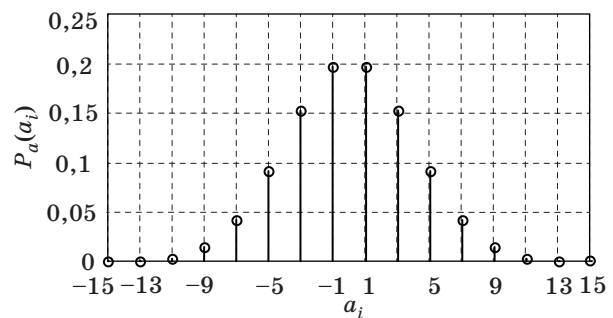
$$P_b(b_k) = C_L^2 q^{\frac{L+b_k}{2}} (1-q)^{\frac{L-b_k}{2}}. \quad (16)$$

Здесь q и $1 - q$ — вероятности передачи символов $+1$ и -1 соответственно. Если $q = 1 - q = 0,5$, то средние значения $m_1(a) = m_1(b) = 0$, дисперсии $\sigma_a^2 = \sigma_b^2 = L$.

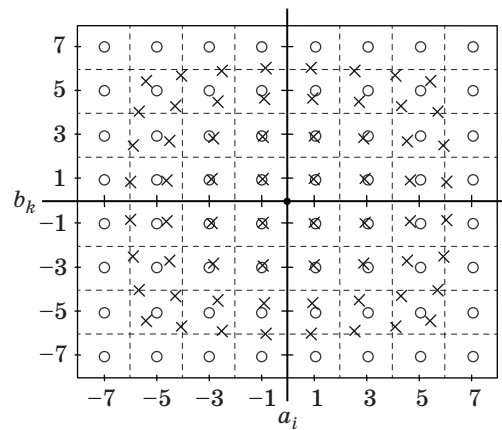
Прохождение группового радиосигнала QAM через нелинейную цепь

Созвездие амплитуд радиосигнала (на примере системы QAM-64 при $L = 7$) показано на рис. 3. Статистические характеристики входных

¹ Обычно $L = N - 1$, так как нулевая (по Уолшу) функция Уолша в качестве переносчика информации не используется.



■ **Рис. 2.** Закон распределения $P_a(a)$ синфазного группового видеосигнала $a(t)$ при $L = 15$, $q = 0,5$



■ **Рис. 3.** Созвездие амплитуд радиосигнала (на примере системы QAM-64). Значения амплитуды чипа: o — на входе НУРС, x — на выходе НУРС

сигналов $a(t)$ и $b(t)$ обоих модуляторов (см. рис. 1) и условия их прохождения через групповой радиотракт идентичны. Поэтому статистические характеристики выходных сигналов обоих демодуляторов, в том числе одномерные плотности распределения, необходимые в дальнейшем для расчета вероятности битовой ошибки, тоже одинаковы. Тогда далее достаточно рассматривать только

один либо два из четырех квадрантов созвездия, например первый, в котором сигналы $a(t)$ и $b(t)$ оба положительны, или первый и второй.

Для дальнейших вычислений применим матричный аппарат. Методика расчета иллюстрируется на примере системы DS-CDMA с QAM-256, построенной на основе ортогонального базиса Радемахера — Уолша размерности $N = 16$, с числом абонентов в каждом стволе $L = 15$. В этом случае сигналы $a(t)$ и $b(t)$ представляют собой дискретные случайные величины a_i и b_k , при-

нимающие в пределах каждого чипа первого квадранта положительные значения из набора 1;3;5;7;9;11;13;15.

Матрица A , показанная ниже, содержит полную совокупность возможных в первом квадранте значений амплитуды $A_{i,k} = \sqrt{a_i^2 + b_k^2}$ радиосигнала на входе нелинейного усилителя радиосигнала НУРС (см. рис. 1). Здесь и далее $i = 0, 1, \dots, 7$ — номер строки, $k = 0, 1, \dots, 7$ — номер столбца матрицы:

$$A = \begin{pmatrix} 1,414 & 3,162 & 5,099 & 7,071 & 9,055 & 11,045 & 13,038 & 15,033 \\ 3,162 & 4,243 & 5,831 & 7,616 & 9,487 & 11,402 & 13,342 & 15,297 \\ 5,099 & 5,831 & 7,071 & 8,602 & 10,296 & 12,083 & 13,928 & 15,811 \\ 7,071 & 7,616 & 8,602 & 9,899 & 11,402 & 13,038 & 14,765 & 16,553 \\ 9,055 & 9,487 & 10,296 & 11,402 & 12,728 & 14,213 & 15,811 & 17,493 \\ 11,045 & 11,402 & 12,083 & 13,038 & 14,213 & 15,556 & 17,029 & 18,601 \\ 13,038 & 13,342 & 13,928 & 14,765 & 15,811 & 17,029 & 18,385 & 19,849 \\ 15,033 & 15,297 & 15,811 & 16,553 & 17,493 & 18,601 & 19,849 & 21,213 \end{pmatrix}.$$

Распределения квадратов чипов a_i^2 и b_k^2 — биномиальные, с количеством возможных значений, равным 8. Так как $a(t)$ и $b(t)$ независимы, распределение амплитуды чипа $A_{i,k}$ тоже биномиальное, но с количеством возможных значений в наборе $8 \times 8 = 64$.

Далее представлена матрица

$$\cos\varphi = \begin{pmatrix} 0,707 & 0,316 & 0,196 & 0,141 & 0,11 & 0,091 & 0,077 & 0,067 \\ 0,949 & 0,707 & 0,514 & 0,394 & 0,316 & 0,263 & 0,225 & 0,196 \\ 0,981 & 0,857 & 0,707 & 0,581 & 0,486 & 0,414 & 0,359 & 0,316 \\ 0,990 & 0,919 & 0,814 & 0,707 & 0,614 & 0,537 & 0,474 & 0,423 \\ 0,994 & 0,949 & 0,874 & 0,789 & 0,707 & 0,633 & 0,569 & 0,514 \\ 0,996 & 0,965 & 0,91 & 0,844 & 0,774 & 0,707 & 0,646 & 0,591 \\ 0,997 & 0,974 & 0,933 & 0,88 & 0,822 & 0,763 & 0,707 & 0,655 \\ 0,998 & 0,981 & 0,949 & 0,906 & 0,857 & 0,806 & 0,756 & 0,707 \end{pmatrix}$$

со значениями $\cos\varphi_{i,k} = a_i/A_{i,k}$ — множителями, необходимыми для выделения сигналов синфазного ствола на приемной стороне. Аналогично определяется набор множителей $\sin\varphi_{i,k} = b_k/A_{i,k}$, образующих матрицу $\sin\varphi = (\cos\varphi)^T$, необходимых для выделения сигналов квадратурного ствола.

Будем считать, что НУРС (обычно это выходной усилитель мощности) является единственным источником нелинейных искажений в радиотракте и что шумом на входе этого усилителя можно пренебречь. Рассмотрим прохождение радиосигнала $S(t)$ через НУРС с нормированной АХ, имеющей вид [1, 3]

$$G(A) = \frac{A}{[1 + (A/s)^{2p}]^{1/(2p)}}, \quad (2)$$

где s — порог ограничения амплитуды радиосигнала; $p \geq 1$ — целочисленный параметр, определяющий степень нелинейности АХ. При $p \geq 10$ АХ (2) становится линейно-ломаной и практически совпадает с АХ идеального [4] ограничителя амплитуды. Если в тракте есть и другие источники нелинейных искажений, например неидеальность модуляционных и (или) демодуляционных характеристик, то они тоже могут быть учтены в выражении типа (2).

Если НУРС — безынерционный, то амплитуды радиосигнала $A_{i,k}$ на входе и искаженные значения $G_{i,k}$ на выходе усилителя имеют одинаковые законы распределения с соответствующими вероятностями

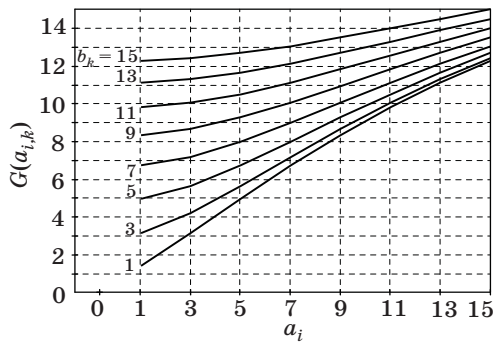
$P_a(a_i) \cdot P_b(b_k)$. Приведенная ниже матрица G содержит совокупность возможных значений $G_{i,k}(A_{i,k})$ при $s = L\sqrt{2}$, $p = 1$:

$$G = \begin{pmatrix} 1,411 & 3,128 & 4,958 & 6,708 & 8,328 & 9,797 & 11,108 & 12,266 \\ 3,128 & 4,16 & 5,622 & 7,168 & 8,66 & 10,043 & 11,294 & 12,408 \\ 4,958 & 5,622 & 6,708 & 7,972 & 9,262 & 10,499 & 11,643 & 12,677 \\ 6,708 & 7,168 & 7,972 & 8,971 & 10,043 & 11,108 & 12,118 & 13,05 \\ 8,328 & 8,66 & 9,262 & 10,043 & 10,914 & 11,808 & 12,677 & 13,496 \\ 9,797 & 10,043 & 10,499 & 11,108 & 11,808 & 12,545 & 13,28 & 13,986 \\ 11,108 & 11,294 & 11,643 & 12,118 & 12,677 & 13,28 & 13,893 & 14,494 \\ 12,266 & 12,408 & 12,677 & 13,05 & 13,496 & 13,986 & 14,494 & 15 \end{pmatrix}.$$

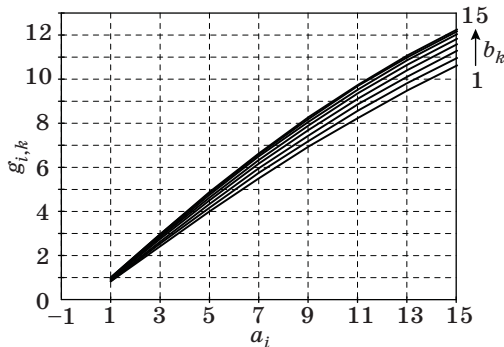
По данным матрицы G построено семейство АХ в первом квадранте $G(a_{i,k}) = G(a_i)$ при $b_k = \text{const}$ (рис. 4).

Матрица g_i , представленная ниже, и матрица $g_k = g_i^T$ суть наборы произведений Адамара $g_{i,k} = G_{i,k} \cos \varphi_{i,k}$ и $g_{k,i} = G_{i,k} \sin \varphi_{i,k}$ элементов соответствующих матриц. Это чипы искаженных квадратурных составляющих группового видеосигнала на двух выходах демодулятора. Законы распределения величин $g_{i,k}$ и $g_{k,i}$ — биномиальные, определяемые выражениями (1а) и (1б):

$$g_i = \begin{pmatrix} 0,998 & 0,989 & 0,972 & 0,949 & 0,92 & 0,887 & 0,852 & 0,816 \\ 2,967 & 2,942 & 2,893 & 2,824 & 2,739 & 2,642 & 2,539 & 2,433 \\ 4,862 & 4,821 & 4,743 & 4,634 & 4,498 & 4,345 & 4,18 & 4,009 \\ 6,641 & 6,588 & 6,487 & 6,343 & 6,166 & 5,964 & 5,745 & 5,519 \\ 8,277 & 8,216 & 8,097 & 7,927 & 7,717 & 7,477 & 7,216 & 6,944 \\ 9,757 & 9,689 & 9,558 & 9,371 & 9,139 & 8,87 & 8,578 & 8,271 \\ 11,075 & 11,004 & 10,867 & 10,67 & 10,423 & 10,138 & 9,824 & 9,492 \\ 12,238 & 12,167 & 12,027 & 11,826 & 11,573 & 11,278 & 10,953 & 10,607 \end{pmatrix}.$$



■ Рис. 4. Семейство АХ НУРС в первом квадранте



■ Рис. 5. Семейство характеристик $g_{i,k} = g(a_i)$ при $b_k = \text{const}$ для синфазной составляющей

По данным матрицы g_i построено семейство графиков $g_{i,k} = g(a_i)$ при $b_k = \text{const}$ для видеосигнала синфазного ствола (рис. 5). Аналогичное семейство кривых $g_{k,i} = g(b_k)$ при $a_i = \text{const}$ для видеосигнала квадратурного ствола можно построить по данным матрицы g_k .

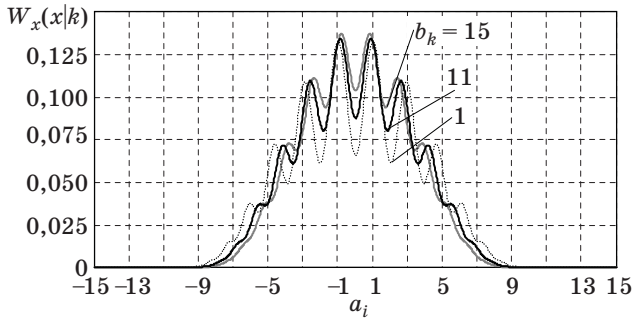
Закон распределения демодулированного сигнала

В приемной части системы (см. рис. 1) на входе нелинейного усилителя видеосигнала НУРС к видеосигналу $g_{i,k}$ добавляется аддитивный гауссовый шум $\varepsilon(t)$ с нулевым средним и дисперсией σ_ε^2 , функция плотности которого

$$W_\varepsilon(\varepsilon) = \frac{1}{\sigma_\varepsilon \sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2\sigma_\varepsilon^2}\right). \quad (3)$$

Плотность распределения смеси $x_{i,k}(t) = g_{i,k}(t) + \varepsilon(t)$ сигнала с шумом на входе НУРС, средняя по всем значениям видеосигнала a_i синфазного ствола, но условная по k , выражается через свертку функций (1а) и (3):

$$W_x(x|k) = \int_{-\infty}^{\infty} \sum_i P_a(a_i) W_\varepsilon(x - g_{i,k}) d\varepsilon =$$



■ Рис. 6. Условная плотность распределения смеси сигнала с шумом на входе НУВС синфазного ствола

$$= \sum_i C_L^2 \frac{L+a_i}{2} q^2 \frac{L-a_i}{2} (1-q)^2 \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp \left\{ -\frac{(x-g_{i,k})^2}{2\sigma_\varepsilon^2} \right\}.$$

Семейство функций $W_x(x|k)$ для двух первых квадрантов представлено на рис. 6. Значения $g_{i,k}$ при $b_k = \text{const}$ связаны со значениями a_i , указанными на оси абсцисс, детерминированным образом, через представленные выше матрицы.

Оценка средней мощности помех на выходе НУВС

Амплитудная характеристика безынерционного НУВС может быть выражена формулой (2), но с заменой переменных A и G на x и y соответственно, а параметров НУРС p и s — на аналогичные параметры НУВС p_B и s_B :

$$y(x) = \frac{x}{\left[1 + (x/s_B)^{2p_B}\right]^{1/(2p_B)}}. \quad (4)$$

Нелинейное преобразование смеси $x(t)$ сигнала с аддитивным шумом в НУВС представим, следуя работам [4, 5], как добавление к полезному сигналу $g_{i,k}$ помехи, состоящей из двух разделяющихся частей: мультипликативной — коррелированной и аддитивной — не коррелированной с этим сигналом. Тогда выходной процесс в пределах отдельно взятого чипа

$$y(t) = \gamma g_{i,k} + \xi(t), \quad (5)$$

где $\gamma g_{i,k}$ — полезная (информативная) часть выходной смеси; $(1 - \gamma)g_{i,k}$ и $\xi(t)$ — соответственно коррелированная и некоррелированная с $g_{i,k}$ составляющие помехи; $0 \leq \gamma \leq 1$ — числовой коэффициент. При этом

$$\overline{\xi^2(t)} = \overline{y^2(t)} - \gamma^2 \overline{g_{i,k}^2}. \quad (6)$$

Для отыскания коэффициента γ умножим на $g_{i,k}$ и усредним по времени в пределах чипа обе части уравнения (5):

$$\overline{g_{i,k}y(t)} = \gamma \overline{g_{i,k}^2} + \overline{\xi(t)g_{i,k}}.$$

Но по условиям задачи $\overline{g_{i,k}y(t)} = \overline{g_{i,k}} \overline{y(t)}$, а $\overline{\xi(t)g_{i,k}} = \overline{\xi(t)} \overline{g_{i,k}} = 0$; тогда

$$\gamma_{i,k} = \overline{y(t)} / \overline{g_{i,k}}. \quad (7)$$

Общие формулы (5)–(7) справедливы и для индексированных по i и k значений переменных y , ξ , γ :

$$y_{i,k}(t) = y[g_{i,k} + \varepsilon(t)]; \quad \overline{\xi_{i,k}^2(t)} = \overline{y_{i,k}^2(t)} - \gamma_{i,k}^2 \overline{g_{i,k}^2};$$

$$\gamma_{i,k} = \overline{y_{i,k}(t)} / \overline{g_{i,k}}.$$

Полагая, что исследуемые процессы эргодичны, заменим усреднение по времени усреднением по множеству реализаций для каждого i и k в отдельности; тогда коэффициент уменьшения полезного сигнала из-за воздействия коррелированной помехи

$$\gamma_{i,k} = \overline{y_{i,k}} / \overline{g_{i,k}}, \quad (8)$$

а дисперсия некоррелированной помехи

$$\sigma_{i,k}^2 = \overline{\xi_{i,k}^2} = \overline{y_{i,k}^2} - \gamma_{i,k}^2 \overline{g_{i,k}^2}. \quad (9)$$

Функция плотности смеси $y_{i,k}$ на выходе НУВС имеет вид [1]

$$W_y(y_{i,k}) = \frac{\exp \left\{ -\frac{1}{2\sigma_{i,k}^2} \left[\frac{y_{i,k}}{2p_B \sqrt{z_{i,k}}} - g_{i,k} \right]^2 \right\}}{2p_B \sqrt{z_{i,k}} z_{i,k} \sigma_{i,k} \sqrt{2\pi}}, \quad (10)$$

где вспомогательная переменная $z_{i,k} = 1 - (y_{i,k}/s_B)^{2p}$. Согласно (8) и (10):

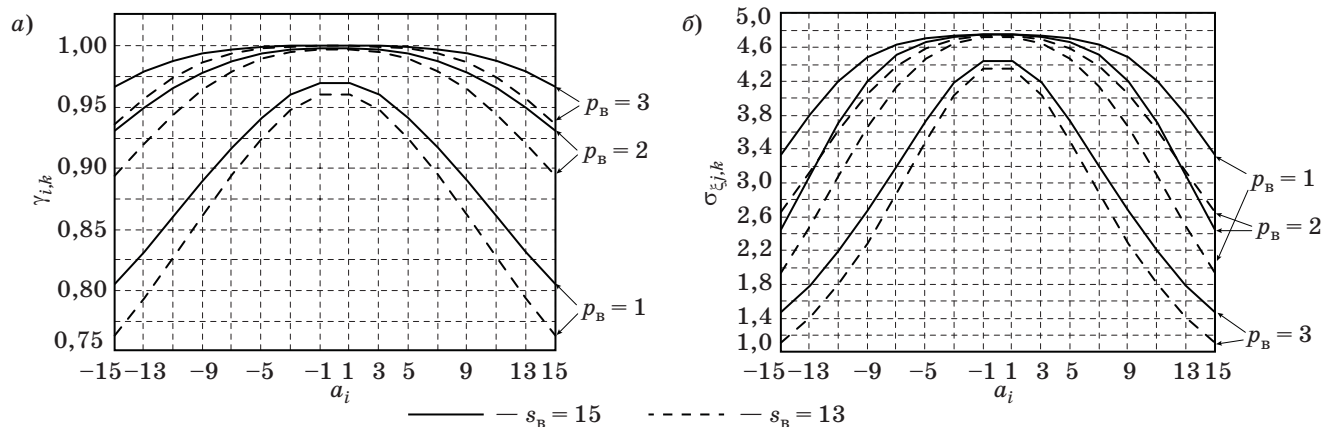
$$\gamma_{i,k} = \frac{\int_{-\infty}^{\infty} y_{i,k} W_y(y_{i,k}) dy_{i,k}}{\overline{g_{i,k}}}. \quad (11)$$

Этот коэффициент зависит не только от i и k , но и от параметров s_B , p_B АХ НУВС, а также, опосредованно, от отношения сигнал/помеха $\rho = 10 \lg(g_{i,k}^2 / \sigma_\varepsilon^2)$ на входе НУВС. Зависимость функции (11) от значений исходного сигнала a_i в первом и втором квадрантах, соответствующих значениям искаженного сигнала $g_{i,k}$ при $k = 15$, $\rho = 10$ дБ, показана на рис. 7, а, а зависимость (9) дисперсии $\overline{\xi_{i,k}^2}$ некоррелированной помехи от a_i — на рис. 7, б.

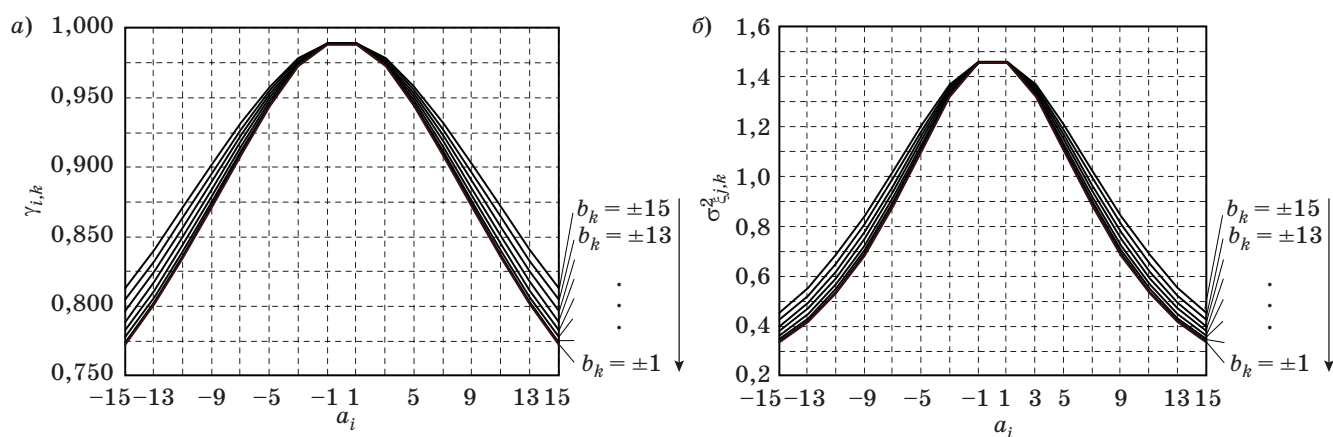
Семейства кривых на рис. 8, а и б иллюстрируют влияние амплитуды b_k чипов квадратурного

ствола на коэффициент $\gamma_{i,k}$ и дисперсию $\overline{\xi_{i,k}^2}$ при $k = \text{var}$. В левой и правой частях обоих графиков это влияние больше, чем в центральной части.

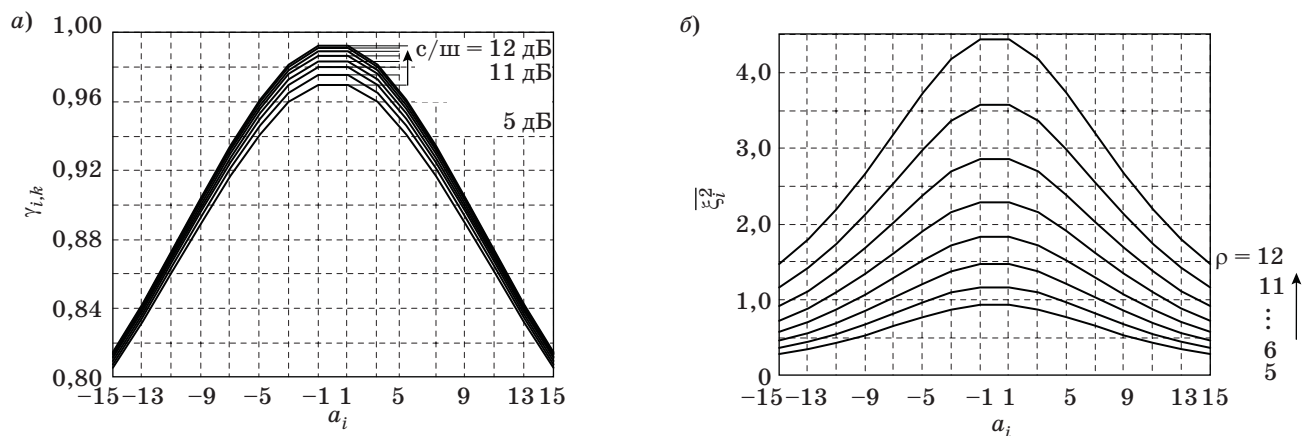
Семейства кривых $\gamma_{i,k}(a_i)$ и $\overline{\xi_{i,k}^2}(a_i)$ при фиксированном $k = 6$ и различных отношениях сигнал/помеха $\rho = 5, 6, \dots, 12$ дБ на входе НУВС показаны



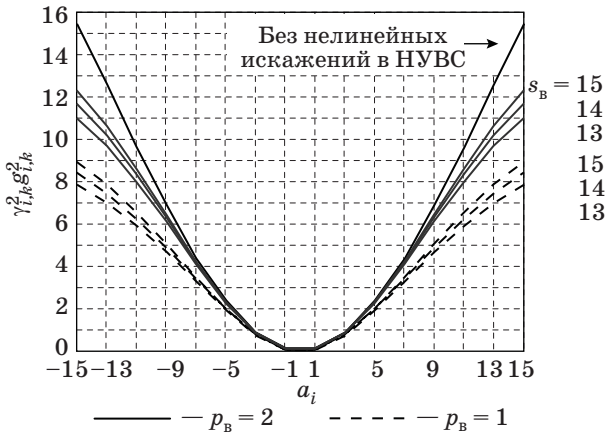
■ Рис. 7. Зависимость коэффициента $\gamma_{i,k}(a)$ и дисперсии $\overline{\xi_{i,k}^2}$ (б) от амплитуды чипа a_i при $k = 15$, $\rho = 10$ дБ



■ Рис. 8. Влияние индекса k на коэффициент $\gamma_{i,k}(a)$ и дисперсию $\overline{\xi_{i,k}^2}$ (б) при $s = \sqrt{2}L$, $p = 1$, $s_B = L$, $p_B = 1$, $\rho = 10$ дБ



■ Рис. 9. Влияние отношения сигнал/шум на коэффициент $\gamma_{i,k}(a)$ и дисперсию $\overline{\xi_{i,k}^2}$ (б) при $s = \sqrt{2}L$, $p = 1$, $s_B = L$, $p_B = 1$



■ Рис. 10. Зависимость мощности полезного сигнала на выходе НУВС от a_i при различных параметрах АХ

на рис. 9, а и б. Чем амплитуда чипа группового сигнала меньше, тем влияние аддитивной помехи больше.

Зависимость мощности $\gamma_{i,k}^2 g_{i,k}^2$ полезного сигнала на выходе НУВС от амплитуды чипа исходного сигнала a_i при $k = 15$ с учетом и без учета нелинейных искажений, вносимых НУРС, представлена на рис. 10.

Дисперсия выходной смеси НУВС согласно (10), (11):

$$\sigma_{y_{i,k}}^2 = \int_{-s_B}^{s_B} y_{i,k}^2 W_y(y_{i,k}) dy_{i,k} = \int_{-s_B}^{s_B} y_{i,k}^2 \frac{\exp\left\{-\frac{1}{2\sigma_{i,k}^2} \left[\frac{y_{i,k}}{2p_B \sqrt{z_{i,k}}} - g_{i,k}\right]^2\right\}}{2p_B \sqrt{z_{i,k} z_{i,k} \sigma_{i,k}^2} \sqrt{2\pi}} dy_{i,k}. \quad (12)$$

Плотность распределения некоррелированной помехи $\xi_{i,k}$ с нулевым средним найдем, подставляя в (10) $\xi_{i,k}$ вместо $y_{i,k}$. При $-s_B < \xi_{i,k} < s_B$

$$W_{\xi}(\xi_{i,k}) = \frac{1/(\sigma_{i,k} \sqrt{2\pi})}{\left[1 - \left(\frac{\xi_{i,k} + \gamma_{i,k} g_{i,k}}{s_B}\right)^{2p_B}\right]^{\frac{1+2p_B}{2p_B}}} \times \exp\left\{-\frac{\left[\frac{\xi_{i,k} + \gamma_{i,k} g_{i,k}}{s_B}\right]^{1/(2p_B)} - g_{i,k}}{2\sigma_{i,k}^2}\right\}^2. \quad (13)$$

В общем случае значение дисперсии некоррелированной помехи на выходе НУВС при помощи выражений (12), (13) может быть найдено методом численного интегрирования.

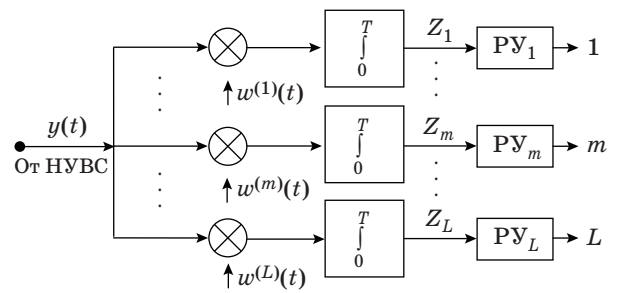
Расчет вероятности битовой ошибки

Демультимплексор каждого ствола (см. рис. 1) содержит L канальных селекторов (рис. 11), осуществляющих в пределах каждого ИО интегрирование с весом $w^{(m)}(t)$ или $w^{(n)}(t)$, и L решающих устройств РУ, принимающих решения о регистрации информационного бита +1 или -1.

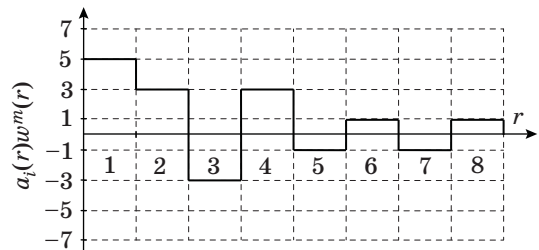
Согласно выражению (5), представим каждый чип процесса $y(t)$ на выходе НУВС в виде суммы полезного группового сигнала $\gamma_{i,k,r} g_{i,k,r}$ и не коррелированной с сигналом помехи $\xi(t)$. Индексы i, k при величинах $\gamma_{i,k,r}$ и $g_{i,k,r}$ связывают эти переменные с теми значениями сигналов a_i и b_k на выходах МП, для которых они рассчитаны выше, а индекс $r = 0, 1, \dots, N-1$ означает порядковый номер чипа в ИО.

Реакция $Z_m(t)$ пары канальный перемножитель — интегратор в канальном селекторе m -го канала (см. рис. 11) на r -й чип группового сигнала синфазного ствола — это сумма сигнальной составляющей $\int_{rT/N}^T \gamma_{i,k,r} g_{i,k,r} w^{(m)}(t) dt$ и помехи $\int_{rT/N}^T \xi_{i,k,r}(t) w^{(m)}(t) dt$. Сумма сигнальных составляющих по всем чипам ИО

$$B_{m,i,k} = \sum_{r=0}^{N-1} \int_{rT/N}^T \gamma_{i,k,r} g_{i,k,r} w^{(m)}(t) dt. \quad (14)$$



■ Рис. 11. Структурная схема ДМП



■ Рис. 12. Выходной сигнал канального перемножителя (без учета нелинейных искажений)

Аналогичный вид имеет выражение для квадратурного ствола. Для примера одна из возможных реализаций сигнала на выходе канального перемножителя показана на рис. 12.

В каждом ИО содержится N чипов, сигналы в которых различным образом сочетаются между собой в зависимости от значений i и k , и нахождение суммарных по всем чипам ИО дисперсий полезного сигнала и помех очень затруднительно из-за многовариантности селектированного полезного сигнала. Однако расчеты можно значительно упростить, если исследовать и учесть комбинаторную структуру этого сигнала. В силу нечетной симметрии АХ (4) в выходном сигнале каждого канального перемножителя в пределах ИО часто возникают пары импульсов противоположной полярности, одинаковых по абсолютной величине. В процессе интегрирования они взаимно компенсируются, причем от взаимного расположения чипов в ИО результат интегрирования не зависит.

Все возможные реализации селектированного сигнала можно найти путем перебора. Машинный эксперимент, проведенный авторами по специально разработанной для этого программе, показывает, что количество возможных различных комбинаций неуравновешенных слагаемых полезного сигнала в ИО невелико, при этом в различных ИО и различных каналах ствола эти комбинации повторяются.

Каждому сигналу $g_{i,k}$ соответствует своя некоррелированная помеха $\xi_{i,k}$. Эти помехи взаимно не компенсируются, поэтому в пределах каждого ИО необходимо учитывать все N слагаемых суммарной помехи η как в «уравновешенных», так и в «неуравновешенных» чипах.

Для упрощения выкладок вернемся к системе QAM-64. Составы и количество возможных комбинаций полезного сигнала B и помехи η , возникающих при любых сочетаниях входных «битов» $+1$ и -1 во всех каналах синфазного ствола с $N = 8$, $L = 7$, приведены в таблице. При любом значении $k = \text{const}$ существует всего три возмож-

Вариант	Комбинации полезного сигнала и помех на выходе канального перемножителя при $k = \text{const}$	Количество комбинаций данного состава
1	$B_1(k) = \gamma_{4,k}g_{4,k} + \gamma_{7,k}g_{7,k}$ $\eta_1(k) = 3\xi_{3,k} + 4\xi_{4,k} + \xi_{7,k}$	50
2	$B_2(k) = \gamma_{5,k}g_{5,k} + \gamma_{6,k}g_{6,k}$ $\eta_2(k) = \xi_{2,k} + 2\xi_{3,k} + 2\xi_{4,k} + 2\xi_{5,k} + \xi_{6,k}$	336
3	$B_3(k) = \gamma_{1,k}g_{1,k} + 4\gamma_{4,k}g_{4,k} + 3\gamma_{5,k}g_{5,k}$ $\eta_3(k) = \xi_{1,k} + 4\xi_{4,k} + 3\xi_{5,k}$	56
Итого		442

ных варианта (индексы 1, 2, 3 при переменных B и η) комбинаций неуравновешенных слагаемых.

Вероятности вариантов равны $\Pi_1 = 50/442 = 0,113$; $\Pi_2 = 336/442 = 0,760$; $\Pi_3 = 56/442 = 0,127$.

Поскольку интегратор представляет собой линейную цепь, полезные сигналы, согласно принципу суперпозиции, суммируются по напряжению, а помехи — по мощности. Поэтому сигнал на выходе интегратора (т. е. на входе РУ) в момент времени T выражается одной из трех комбинаций $B_1(k)$, $B_2(k)$ или $B_3(k)$, а помеха — соответствующей комбинацией $\eta_1(k)$, $\eta_2(k)$ или $\eta_3(k)$ с нулевым средним и дисперсией

$$\begin{aligned}\sigma_{\eta_1}^2(k) &= 3\sigma_{\xi_{3,k}}^2 + 4\sigma_{\xi_{4,k}}^2 + \sigma_{\xi_{7,k}}^2; \\ \sigma_{\eta_2}^2(k) &= \sigma_{\xi_{2,k}}^2 + 2(\sigma_{\xi_{3,k}}^2 + \sigma_{\xi_{4,k}}^2 + \sigma_{\xi_{5,k}}^2) + \sigma_{\xi_{6,k}}^2; \\ \sigma_{\eta_3}^2(k) &= \sigma_{\xi_{1,k}}^2 + 4\sigma_{\xi_{4,k}}^2 + 3\sigma_{\xi_{5,k}}^2.\end{aligned}$$

Распределение суммы взаимно независимых случайных величин $\xi_{i,k}$ на входе РУ далее аппроксимируем нормальным законом с помощью ряда Грама — Шарлье типа А [6]. Если ограничиться первыми членами ряда, то условная по k плотность случайной величины $\eta_1(k)$ равна (ниже для сокращения записи громоздких формул аргумент k в правой части опущен)

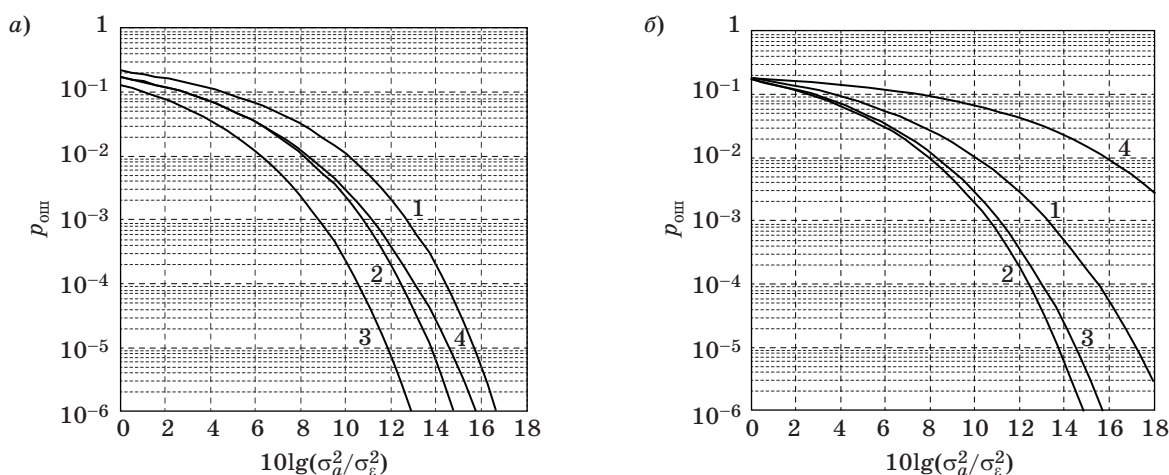
$$\begin{aligned}w_{\eta_1}(\eta_1|k) &= \frac{\exp(-\lambda_1^2/2)}{\sigma_{\eta_1}\sqrt{2\pi}} \times \\ &\times \left[1 + \mu_3\{\eta_1\}(\lambda_1^3 - 3\lambda_1) / (\sigma_{\eta_1}^3 \cdot 3!) + \right. \\ &\left. + \mu_4\{\eta_1\}(\lambda_1^4 - 6\lambda_1^2 + 3) / (\sigma_{\eta_1}^4 \cdot 4!) \right], \quad (15)\end{aligned}$$

где $\lambda_1 = \eta_1(k)/\sigma_{\eta_1}(k)$; μ_3 и μ_4 — третий и четвертый центральные моменты помехи η_1 , определяемые [7] через моменты ее слагаемых:

$$\begin{aligned}\mu_3\{\eta_1\} &= \sum_i \mu_3\{\xi_{i,k}\}; \\ \mu_4\{\eta_1\} &= \sum_i \mu_4\{\xi_{i,k}\} + 6 \sum_{i < r} \sigma_{\xi_{i,k}}^2 \sigma_{\xi_{r,k}}^2,\end{aligned}$$

где $r = 3; 3; 3; 4; 4; 4; 7$. Для функций плотности $W_{\eta_2}(\eta_2|k)$ и $W_{\eta_3}(\eta_3|k)$ формулы аналогичны, но наборы индексов r при вычислении моментов помех другие: $r = 2; 3; 3; 4; 4; 5; 5; 6$ для $\mu_4\{\eta_2(k)\}$ и $r = 1; 4; 4; 4; 4; 5; 5; 5$ для $\mu_4\{\eta_3(k)\}$. В каждом из трех случаев суммируются только слагаемые с теми индексами i , которые указаны в соответствующей строке таблицы.

Для того чтобы функции плотности, приближенно представляемые в виде (15), не оказались отрицательными, должно выполняться [6] условие $(\mu_3\{\eta\}/\sigma_{\eta}^3) < 0,5$. Проверка показывает, что



■ Рис. 13. Вероятность битовой ошибки в абонентском канале системы DS-CDMA с QAM: а — условные вероятности ошибки $p_{ош1}, \dots, p_{ош3}$; б — средняя $p_{ош}$ при разных методиках расчета

оно выполняется для всех η во всем практически возможном диапазоне параметров p, s, p_b, s_b .

Ошибка в РУ с нулевым порогом решения происходит, если при сигнале $B_1(k) > 0$ в момент решения T помеха $\eta_1(k) < -B_1(k)$ или при сигнале $B_1(k) < 0$ помеха $\eta_1(k) > |B_1(k)|$. Условная по k вероятность ошибки при $q = 1 - q = 0,5$

$$p_{ош1}(k) = 0,5 \int_{-B_1(k)}^{\infty} W_{\eta_1}(\eta_1|k) d\eta_1 + 0,5 \int_{B_1(k)}^{\infty} W_{\eta_1}(\eta_1|k) d\eta_1.$$

Условные вероятности $p_{ош2}(k)$ и $p_{ош3}(k)$ вычисляются аналогично. Графические зависимости усредненных по всем k с учетом (1а) вероятностей $p_{ош1}, p_{ош2}, p_{ош3}$ от отношения $10\lg(\sigma_a^2/\sigma_\epsilon^2)$ при $L = N - 1 = 7, \sigma_a^2 = L, s = L\sqrt{2}, p = 1, s_b = 1, p_b = 1$ представлены на рис. 13, а (кривые 1, 2, 3 соответственно). Там же (кривая 4) показан график результирующей вероятности битовой ошибки $p_{ош}$, средней по всем k и по трем комбинаторным вариантам (см. таблицу):

$$p_{ош} = P_1 p_{ош1} + P_2 p_{ош2} + P_3 p_{ош3}.$$

Для сигналов других квадрантов обоих стволов, ввиду их статистической идентичности, промежуточные выкладки и конечные результаты аналогичны.

Для сравнения на рис. 13, б представлены графики зависимости вероятности $p_{ош}$ битовой ошибки на выходе канала от отношения $10\lg(\sigma_a^2/\sigma_\epsilon^2)$, рассчитанные для системы передачи CDMA с QAM-64 с исходными данными, приведенными выше, при следующих условиях:

— с учетом отрицательно действующего на помехоустойчивость фактора — нелинейных иска-

жений в НУРС и НУВС; без учета влияния положительного фактора — комбинаторных свойств группового сигнала Уолша (кривая 1);

— с учетом нелинейных искажений и комбинаторных свойств группового сигнала Уолша, но с заменой реального распределения помех (13) на входе канального РУ гауссовым распределением (кривая 2); в результате такой замены расчетная помехоустойчивость канала связи завышается по сравнению с реальной;

— с учетом нелинейных искажений и комбинаторных свойств группового сигнала Уолша по предложенной в статье методике, с аппроксимацией (15) распределения помехи (13) на входе канального РУ при помощи модифицированных полиномов Эрмита и рядов Грама — Шарлье (кривая 3). Эта кривая, промежуточная между кривыми 1 и 2 и повторяющая кривую 4 (см. рис. 13, а), наиболее близка к реальности.

На рис. 13, б представлен также известный (например, [8]) график вероятности ошибки в системе с QAM-64 без CDMA, с отдельной регистрацией каждого чипа (кривая 4). При равновероятности и независимости всех чипов скорость передачи информации в каждом стволе такой системы составляет 3 бита/чип. График 4 завышает реальную помехоустойчивость системы и, в частности, не учитывает, что ошибка при регистрации одного чипа приводит к искажению от 1 до 3 битов входящего цифрового потока.

Заключение

Представленное исследование реальной помехоустойчивости системы DS-CDMA с QAM дает возможность оценить вероятность битовой ошибки в абонентском канале с учетом нелинейности АХ групповых видео- и радиотрактов и задать требуемое отношение сигнал/шум в различных

частях системы передачи. Компьютерный эксперимент по оценке статистических и комбинаторных свойств группового сигнала Уолша, поставленный авторами, существенно уточняет и одновременно упрощает многофакторные расчеты. Методика анализа, изложенная на примерах систем CDMA с базисом ортогональности $N = 8$

(QAM-64) и (частично) базисом $N = 16$ (QAM-256), полностью применима и при других базисах.

Результаты работы могут быть использованы при теоретических исследованиях и в процессах технической разработки, производства, частотно-территориального планирования высокоскоростных линий передачи.

Литература

1. Гуревич В. Э., Егоров С. Г. Нелинейные искажения группового сигнала в радиосистеме абонентского доступа // *Инфокоммуникационные технологии*. 2012. Т. 10. № 3. С. 79–82.
2. Егоров С. Г. Анализ реальной помехоустойчивости группового тракта радиосистемы абонентского доступа // *Инфокоммуникационные технологии*. 2012. Т. 10. № 4. С. 59–62.
3. Rapp C. Effects of HPA-Nonlinearity on a 4-DPSK/OFDM-Signal for a Digital Sound Broadcasting System // *Proc. Second European Conf. on Satellite Com-*

- munications*, Liege, Belgium, Oct. 22–24, 1991, pp. 179–184.
4. Левин Б. Р. Теоретические основы статистической радиотехники. Кн. 1. — М.: Сов. радио, 1984. — 550 с.
5. Харкевич А. А. Борьба с помехами. — М.: Либроком, 2013. — 280 с.
6. Kendall M., Stuart A. *Distribution Theory (Kendall's The Advanced Theory of Statistics, 1)*. 2nd ed. — London, Charles Griffin & Co., 1962. — 676 p.
7. Вентцель Е. С. Теория вероятностей. — М.: Высш. шк., 1969. — 576 с.
8. Bellamy J. *Digital Telephony*. — N. Y.: Wiley, 2000. — 643 p.

UDC 621.396.49

doi:10.15217/issn1684-8853.2015.5.106

Performance Analysis of Nonlinear Distortions of RF Signal in Wireless CDMA Downlink with Quadrature Amplitude Modulation

Gurevich V. E.^a, PhD, Tech., Professor, gurve23@mail.ru

Egorov S. G.^a, Researcher, Assistant, sgegorov@gmail.ru

^aThe Bonch-Bruевич Saint-Petersburg State University of Telecommunications, 22-1, Bolshievikov St., 193232, Saint-Petersburg, Russian Federation

Purpose: In this paper, we consider digital direct-sequence spread-spectrum multiple-access communication systems with CDMA technology and QAM modulation. Non-linear distortion of a grouped signal in these systems lead to a higher probability of digital errors and worse quality of individual subscriber's channels. The aim of this research is to find out how the digital error probability depends on the non-linear distortion of grouped video and radio signals. We need to know this dependence to substantiate the demands on the amplitude characteristics of the amplifiers in multiple-duct video and radio channels. **Methods:** A real grouped signal is represented as a sum of three components: the non-distorted part of the signal, the correlated multiplicative distortion and uncorrelated additive distortion. A computer model of the transmitted signal is built and used, taking into account, unlike the conventional methods, not only statistical but also combinatorial properties of a complex grouped signals with code division. An admissible error of estimated noise immunity is obtained by the approximation of real distribution of the signal and noise mixture, using Hermite polynomials and Gram-Charlier series. **Results:** It is shown that the usual Gaussian representation of the noise and the linearization of non-linear elements characteristics lead to overestimating the noise immunity. Analyzing the combinatorics of the grouped signal has helped to estimate both separate and combined impact of each parameter of nonlinear channels upon the real noise immunity of the system. The laws of distribution of the mixture of Rademacher – Walsh functions with a noise in a nonlinear channel are represented analytically and graphically, along with the calculated error probabilities. **Practical relevance:** Our recommendations about technical demands on grouped channels for systems with code division can improve the noise stability. They can be used in theoretical research, technical development, production and application of new high-speed digital transmission systems.

Keywords — DS-CDMA Wireless Access, Orthogonal Rademacher – Walsh Functions, Quadrature Amplitude Modulation, Nonlinear Amplifiers, Radio Frequency Power Amplifiers, Multiplexed Signal Characteristics, Partial Cross-Correlation, Correlated Distortions, Uncorrelated Distortions, Signal-to-Noise Ratio, Bit Error Rate.

References

1. Gurevich V. E., Egorov S. G. Nonlinear Distortions of Multiplexed Signal in Wireless Local Loop System. *Infokommunikatsionnye tekhnologii*, 2012, vol. 10, no. 3, pp. 79–82 (In Russian).
2. Egorov S. G. Error Rate Analysis of Multiplexed Signal in Imperfect Wireless Local Loop System. *Infokommunikatsionnye tekhnologii*, 2012, vol. 10, no. 4, pp. 59–62 (In Russian).
3. Rapp C. Effects of HPA-Nonlinearity on a 4-DPSK/OFDM-Signal for a Digital Sound Broadcasting System. *Proc. Second European Conf. on Satellite Communications*, Liege, Belgium, October 22–24, 1991, pp. 179–184.

4. Levin B. R. *Teoreticheskie osnovy statisticheskoi radiotekhniki* [Theoretical Foundations of the Statistical Radio Engineering]. Moscow, Sovetskoe radio Publ., 1974. 550 p. (In Russian).
5. Kharkevich A. A. *Bor'ba s pomexami* [Noise control]. Moscow, Librokom Publ., 2013. 280 p. (In Russian).
6. Kendall M., Stuart A. *Distribution Theory (Kendall's the Advanced Theory of Statistics, 1)*. 2nd ed., London, Charles Griffin & Co., 1962. 676 p.
7. Venttsel' E. S. *Teoriia veroiatnostei* [Probability Theory]. Moscow, Vysshiaia shkola Publ., 1969. 576 p. (In Russian).
8. Bellamy J. *Digital Telephony*. New York, Wiley, 2000. 643 p.

МАСКИРОВАНИЕ ЦИФРОВОЙ ВИЗУАЛЬНОЙ ИНФОРМАЦИИ: ТЕРМИН И ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

А. А. Востриков^а, канд. техн. наук, доцент

М. Б. Сергеев^а, доктор техн. наук, профессор

М. Ю. Литвинов^б, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

^бМосковский государственный технический университет им. Н. Э. Баумана, Москва, РФ

Постановка проблемы: в настоящее время для идентификации симметричных преобразований цифровой визуальной информации в целях ее сокрытия от несанкционированного ознакомления существует ряд терминов. Значительные расхождения в реализации преобразований, применение в них разных математических методов и связанных с этим требований к эффективности сокрытия приводят к необходимости уточнения соответствия терминов выполняемым процедурам. Ставится цель критически оценить термины «маскирование» и «демаскирование», рассмотреть альтернативные варианты терминологического обозначения преобразований, выполняемых для защиты цифровых изображений с малым временем актуальности, и закрепить наиболее подходящий вариант. **Результаты:** сопоставительный анализ терминологии, близкой по значению к описанию содержания процедур преобразования цифровой визуальной информации, показал, что в настоящее время отсутствуют сколь-нибудь более близкие термины, чем «маскирование» и «демаскирование», с определением, способным указать точное место рассматриваемых процедур среди способов защиты информации. **Практическая значимость:** показано, что термин «маскирование» и связанный с ним термин «демаскирование» наиболее близки к описанию сути преобразований цифровой визуальной видеoinформации с малым временем актуальности в целях ее сокрытия от несанкционированного ознакомления. Сформулированы основные определения применительно к матричным преобразованиям цифровой визуальной информации с использованием квазиортогональных базисов, приведены матричные соотношения для выполнения «маскирования» и «демаскирования».

Ключевые слова — защита визуальной информации, защита изображений, маскирование, демаскирование, квазиортогональные базисы, квазиортогональные матрицы, криптографические примитивы.

Введение

Термины и терминология в целом составляют лингвистический инструментарий сообществ, работающих в определенной области знаний. В научной среде термины дополняют естественный язык и, таким образом, расширяют из года в год словарное множество в соответствующем направлении исследований. Появление новых терминов в большинстве случаев вызвано либо необходимостью замены неудобных трудновоспроизводимых или просто длинных языковых конструкций, либо открытием новых объектов, для которых ранее в языке определение не существовало.

Характерно то, что зачастую термины в их написании и произношении заимствуются из совершенно других областей человеческой деятельности, увеличивая омонимию или полисемию какого-либо слова. Это, как правило, естественный процесс, являющийся следствием стремления исследователей к удобству их восприятия и привычности звучания, хотя нередко причиной является просто отсутствие в языке более близких по значению понятий.

Предмет рассмотрения настоящей работы — термин «маскирование» — довольно обширен и уже длительное время используется в различных областях человеческой деятельности, таких как биология, военное дело, химия, психология,

технология управления базами данных (БД), обработка фотографического изображения, область защиты аналоговых сигналов, цифровая обработка изображений.

И уже более восьми лет им пользуются в активно развивающейся научно-технической области — защита от несанкционированного доступа к цифровым изображениям с малым периодом актуальности.

Впервые такая терминология в отношении цифровых изображений была введена в работе [1], а затем более обстоятельно — в работе [2]. Предложенный вариант термина позволял авторам дистанцироваться от традиционных для защиты информации криптографических преобразований, которые, как было показано, в применении к цифровой визуальной информации и при реализации в распределенных встраиваемых системах имеют существенные недостатки [3].

В дальнейшем данный термин был подхвачен и используется до сих пор авторами работ, активно развивающими теоретическую и прикладную стороны применения нового множества недавно открытых вещественных квазиортогональных М-матриц [4].

Настоящая работа ставит своей целью критически оценить применение термина «маскирование» в прикладных задачах защиты цифровой визуальной информации от несанкционирован-

ного доступа, очертить границы применения и зафиксировать этот термин для терминологического описания процедур в указанных задачах.

Области применения термина «маскирование»

Маскирование в биологии. Основополагающие принципы маскирования (маскировки) заложены самой природой (мимикрия). Ряд видов, которые являются беззащитными перед некоторыми хищниками, имитируют другой вид, который избегается хищниками ввиду своей несъедобности или наличия специфических средств защиты. Соответственно, в биологии под маскированием понимается изменение внешнего вида животных, преимущественно окраски, в соответствии с условиями окружающей среды [5].

Маскирование в военном деле — комплекс мероприятий, направленный на введение противника в заблуждение относительно наличия и расположения войск, военных объектов, их состояния и планов командования [6].

Маскирование в аналитической химии — процесс, позволяющий ускорить и упростить анализ смеси при снижении влияния отдельных ее компонентов на разделение, обнаружение или определение веществ без отделения маскируемого (мешающего) соединения от исследуемой смеси [7].

Маскирование в психологии — общий термин в исследованиях восприятия, обозначающий любой процесс, посредством которого различаемый (распознаваемый) стимул становится трудно или совсем не различимым (нераспознаваемым) из-за влияния второго стимула (маскировщика) во временной или пространственной близости к первому [8].

Маскирование изображений в психологии — процесс предъявления тестового зрительного стимула на фоне маскирующего структурированного стимула (изображения), в результате чего ухудшается обнаружение и опознание тестового стимула [9].

Маскирование данных (datamasking) — процесс идентификации конфиденциальных данных и наложения на них «защитной маски», которая сохраняет их «неприкосновенность» в БД, не нарушая при этом функциональной целостности приложения, использующего эти данные [10].

При маскировании данных заменяется конфиденциальная информация в БД, например номера кредитных карт или номера социального страхования на реалистичные значения, что позволяет безопасно использовать эти данные для разработки и тестирования или обмениваться ими с внешними партнерами. Другими словами, маскирование данных — процесс обезличива-

ния конфиденциальной информации, хранящейся в БД [11].

Маскирование видеосигнала — способ защиты видеоизображения, представленного в аналоговой форме, как правило, путем перестановки участков сигнала, содержащих информацию о различных фрагментах изображения (чаще всего — строк). В англоязычном варианте устройства, осуществляющие маскирование видеосигнала, — маскираторы — известны под названием скремблеры (от англ. scramble — перемешивать, перепутывать) [12].

Маскирование речи — способ изменения аналогового или цифрового звукового сигнала в целях искажения характеристик голоса или сокрытия информационной составляющей речевого сообщения [13].

Нерезкое маскирование в области обработки фотографического изображения — технологический прием, позволяющий добиться эффекта ощущения большей резкости за счет усиления контраста тональных переходов [14].

Маскирование слоев в цифровой обработке изображений — способ создания композиций, предназначенный для объединения фотографий в единое изображение, а также для проведения локальных корректировок цветности и тона [15].

Маскирование цифровой визуальной информации. Наконец, в уже упомянутых работах [1, 2] вводится понятие маскирования цифровой визуальной информации как процесса преобразования изображения в кадре к шумоподобному виду. Для маскирования в этом случае используются криптографические примитивы и матричные преобразования.

Таким образом, общее толкование термина «маскирование» для различных областей его применения распадается на две группы:

1) маскирование как действие по сокрытию (защите, деактивации, независимой обработке) определенных частей чего-то целого (аналитическая химия, психология, фотография, цифровая обработка изображений);

2) маскирование как действие с целью обеспечить невозможность обнаружения, идентификации объектов и их частей (биология, военное дело, системы управления БД, защита визуальной информации).

Этимология термина

Термин «маскирование», разумеется, происходит от существительного «маска», имеющего прототипическое значение: «то, что скрывает что-либо» [16]. Из школьного этимологического словаря известно, что русское слово «маска» заимствовано в начале XVIII в. из французского языка (слово «masque») и итальянского

аналога «maschera», восходящего к арабскому слову «maskhara», означающего «маска», образованному в свою очередь от слова «насмешка».

В соответствии со словарем современного русского языка Ефремовой [17] слово «маскирование» означает процесс действия по глаголам: маскировать, маскироваться. К слову же «маскировать» приводятся следующие значения:

— одевать в маскарадный или не свойственный кому-либо костюм или надевать маску;

— переносное: делать что-либо незаметным, невидимым, скрывать что-либо посредством чего-либо показного, притворного;

— скрывать от противника с помощью разных приемов.

Итак, в прототипическом значении «маска» — это то, что скрывает что-либо. При этом в ряде перечисленных выше областей термин «маскирование» применяется не в значении спрятать целевой объект целиком, а выборочно защитить или скрыть лишь определенные его части.

В применении же к защите цифровой визуальной информации маскирование понимается как полное сокрытие целевой информации от ознакомления при отсутствии «ключа» у пытающегося открыть защиту. «Ключ» же предоставляется только доверенному кругу и скрывается от кого-либо еще. Таким образом, это определение полностью совпадает с одним из прямых значений слова из словаря Ефремовой: «Скрывать от противника с помощью разных приемов».

Специфика определения и применения термина в области защиты цифровой визуальной информации

Одновременно с очевидной схожестью определения термина в целевой области с уже существующими определениями бросается в глаза и некоторое отличие. Оно заключается в том, что во всех иных значениях при маскировании не происходит изменение целевого объекта или его частей. А в применении к рассматриваемой области целевой объект (исходное изображение) подвергается преобразованию, т. е. изменяется сам. Данное отличие является, очевидно, наиболее существенным от устоявшихся уже вариантов применения термина, хотя цель сокрытия при этом сохраняется и выполняется.

Кроме этого, следует отдельное внимание уделить ударению в рассматриваемом слове. Дело в том, что в интересующем нас контексте «де-факто» закрепилось произношение: «маскИровать» и «маскИрование». Хотя, строго говоря, существующие словари не поддерживают, например, в слове «маскИровать» ударение на второй слог. Более того, часто такое ударение специально указывается как неправильное и утвержда-

ется, что верен лишь вариант «маскировАть» во всех его известных значениях.

В русском языке есть слова, в которых смысл напрямую зависит от их ударения. Типичный пример — слово «броня»: если ударение падает на первый слог, слово означает «закрепление», если же на последний — «защитную оболочку». То же относится к соответствующим глаголам: «брОня» — «бронИровать» (официально закреплять что-либо за кем-либо) и «бронЯ» — «бронировАть» (покрывать бронёй). Однако в этом случае фактически речь идет о двух словах с разным значением и происхождением. А варьирование ударения в слове «маскирование» смысл не меняет.

Тем не менее известно, что есть тенденции общего плана, которые затрагивают акцентные системы разных классов слов в русском языке. Такой общей тенденцией в ударении, по признанию многих лингвистов, считается так называемая тенденция к ритмическому равновесию, которая заключается в смещении ударения к центру слова. Процесс смещения ударения в многосложных словах к центру слова оправдан с практической точки зрения, так как в этом случае соблюдается более равномерная смена ударных и неударных слогов, что более удобно для произнесения [18]. Учитывая вместе с этим упомянутые выше специфические отличия маскирования как защиты визуальной информации от прочих вариантов толкования, можно констатировать, что в данном случае изменение ударной гласной позволяет термину подчеркнуть особенный вариант его применения.

Альтернативные варианты терминов

С учетом относительной новизны применения рассматриваемого термина целесообразно также рассмотреть возможные варианты обозначения соответствующего действия и образованных от них существительных. При этом близость смыслового определения следует соотносить с благозвучностью терминов, в противном случае не будет гарантии, что термин приживется. В таблице приведены наиболее близкие по смыслу к рассматриваемому действию глаголы и существительные.

Как видно из таблицы, при достаточном многообразии близких по значению слов не удастся найти термин, который можно было бы считать более подходящим для применения.

Выбор слова «маскирование» выглядит удачной находкой, поскольку позволяет указать место, занимаемое соответствующими преобразованиями в ряду методов защиты цифровой визуальной информации от несанкционированного доступа. Маскирование визуальной информации

■ Альтернативные варианты термина

Глагол	Существительное	Оценка
Скрывать	Соккрытие	Просторечное, «бытовое» звучание
Укрывать	Укрытие	
Прятать	Прятки	
Защищать	Защита	Общее понятие, требующее конкретизации с помощью длинной конструкции
Ограничивать	Ограничение	
Преобразовывать	Преобразование	
Камуфлировать	Камуфлирование/Камуфляж	Не соответствуют в точности целевому определению
Подменять	Подмена	
Шифровать	Шифрование/Шифрация	Предполагает полноценные криптографические преобразования, применение которых для целевого определения существенно ограничено [3]

не предназначено для разграничения доступа в процессе длительного хранения, как криптография, и, следовательно, ориентировано на данные с малым временем актуальности. Кроме этого, предмет защиты — изображение — по своей сути всегда является массивом со значительной степенью избыточности, что при шифровании не гарантирует полное устранение информативности для третьей стороны. В эталонном случае воспроизводимая визуальная информация распадается до белого шума. Но данный критерий труднодостижим, поэтому процедура маскирования предназначена для преобразования изображения в шум с неясным распределением, где не прослеживаются контуры скрываемого объекта или субъекта, даже в случае его перемещения в последовательности кадров видеоизображения.

Основные определения

На основании приведенных в настоящей работе сведений о процедуре маскирования цифровой визуальной информации, а также информации из опубликованных работ в целевом направлении исследований [1, 2, 3, 19–21] сформулируем основные определения терминов.

Маскирование — процесс преобразования цифровой визуальной информации с малым периодом актуальности к шумоподобному виду в целях ее защиты от несанкционированного ознакомления.

После выполнения маскирования полученный массив информации называется *маскированной цифровой визуальной информацией* или *маскированным изображением*.

Демаскирование — процесс обратного преобразования маскированной цифровой визуальной информации путем применения операций, являющихся обратными к маскирующим операциям, в целях восстановления исходного изображения.

Матричное маскирование и демаскирование — выполнение соответствующей процедуры прямого или обратного преобразования с применением базиса квазиортогональных матриц.

Для матричного маскирования матрицы должны выбираться так, чтобы они удовлетворяли следующим условиям [22]: квадратная матрица A порядка n с элементами $|a_{ij}| \leq 1$, определенная над полем вещественных чисел и удовлетворяющая квадратичному уравнению связи $A^T A = \omega(n)I$, где $\omega(n)$ — некоторая весовая функция, определяющая тип матрицы, а I — единичная матрица.

Криптографическое маскирование и демаскирование — выполнение соответствующей процедуры прямого и обратного преобразований с применением элементов криптографических методов.

Примечание: Под цифровой визуальной информацией понимаются как сохраненные или передаваемые статические изображения (данные, представляемые в виде двумерного массива значений яркости или цвета минимальной неделимой части такого изображения), так и видеоизображения, понимаемые как изображения, последовательно сменяющие друг друга во времени.

Криптографическое маскирование и демаскирование

При данном виде маскирования цифровой визуальной информации применяются криптографические методы или их элементы в совокупности с какими-либо еще преобразованиями.

Авторы методов криптографического маскирования сталкиваются с необходимостью вырабатывать специализированные подходы, позволяющие разрушить в маскированном изображении низкочастотные пространственные составляющие. Это связано с естественным свойством изображения как вида информации — избыточ-

ностью. Очевидно, что особенно остро такая проблема стоит при работе с высококонтрастными изображениями, содержащими общие планы или крупные объекты.

Кроме того, приходится сокращать вычислительную емкость собственно криптографических методов, нацеленных на длительную (в идеале — навсегда) защиту от несанкционированного доступа, обойти которую можно лишь путем подбора ключевой информации. Сокращение вычислительной емкости при маскировании изображений с малым временем актуальности по определению не требуется, а объем информации в то же время чрезвычайно велик.

Например, авторы работ [2, 3] предлагают изменить модификацию блокового преобразования (например, в виде простейших операций замены и перестановки, аналогичных ГОСТ 28147-89) с некоторым изменяющимся параметром K , что позволяет обеспечить эффективное искажение структуры изображения и тем самым предотвратить наиболее распространенные методы атак. Само преобразование выполняется с использованием схемы сетей Фейстейла с не менее чем восемью раундами, устойчивой ко всем известным в настоящее время атакам. Параметр K меняется для каждого нового блока изображения. Для обеспечения быстрой синхронизации, например при потере пакетов в коммуникационных системах, начальное значение параметра K_{i0} для пакета с номером i может быть установлено следующим образом: $K_{i0} = F(K_0, i)$.

В качестве функции $F(*,*)$ может быть выбрано простое арифметическое сложение 256-битных чисел с игнорированием разряда переполнения.

Преимуществом криптографического маскирования является более высокая устойчивость к атакам по сравнению с другим видом маскирования, недостатком — настолько же более высокая вычислительная ресурсоемкость.

Матричное маскирование и демаскирование

При матричном маскировании цифровой визуальной информации используется матричная арифметика и не используются криптографические подходы. Прикладной предпосылкой для защиты интенсивных потоков с помощью матричной арифметики сегодня является широкое применение интегральных схем цифровых сигнальных процессоров и программируемой логики в качестве вычислительной основы распределенных, встраиваемых и портативных устройств. Интегральные схемы данных классов обладают аппаратными модулями, ускоряющими операцию свертки и, соответственно, скалярного произведения векторов. Как результат, матричное

умножение выполняется на аппаратном уровне, что определяет наилучшую производительность реализации вычислителя при одной и той же полупроводниковой технологии производства интегральных схем.

Использование матричных операций для маскирования цифровой визуальной информации первыми предложили И. Л. Ерош и М. Б. Сергеев [1]. Недостатком предложенного метода являлось то, что сама предложенная матрица — оператор преобразования — очень ресурсоемка для вычисления, а количество таких матриц крайне ограничено. Данного недостатка лишены новые уникальные ортогональные базисы, включающие матрицы Мерсенна, Эйлера, Ферма, Мерсенна — Уолша и др. [23–26].

Такие матрицы хотя и включают в себя иррациональные значения, но обладают одним замечательным свойством: количество различных значений элементов (уровней) в общем случае исчисляется единицами [27]. Это свойство существенным образом способствует выбору в их пользу для прикладного применения, поскольку при вычислении произведения матриц малое количество возможных вариантов множителя позволяет исключить операцию умножения — она заменяется табличной выборкой, что увеличивает производительность вычислителя, сокращает необходимые аппаратные ресурсы.

В общем случае действия по матричному маскированию и демаскированию выполняются следующим образом. Подлежащее преобразованию изображение представляется в виде двумерной матрицы, элементы которой являются численными значениями яркости его пикселей. Суть метода матричного маскирования заключается в умножении сегмента исходного изображения (двумерного набора пикселей) на маскирующую матрицу в виде $S = P \cdot M$, где S — маскированное изображение; P — сегмент исходного изображения; M — матрица преобразующего оператора (маскирующая матрица).

В варианте, дающем лучшие результаты, маскирование осуществляется как $S = M \cdot P \cdot M^T$, где M^T — транспонированная маскирующая матрица.

Восстановление маскированного изображения осуществляется обратным преобразованием, т. е. умножением матрицы S на матрицы, обратные M и M^T , в виде $S = M^T \cdot P \cdot M$.

К настоящему времени опубликовано значительное количество работ, описывающих результаты исследований относительно поиска [28, 29] и применения найденных вещественных квази-ортогональных матриц для защиты визуальной информации [19–22, 30, 31]. Кроме того, осуществляется исследовательская работа для проверки идеи о возможности одновременного решения за-

дачи защиты от несанкционированного ознакомления и задачи сжатия визуальной информации.

Действительно, применение таких матриц как базиса ортогональных преобразований позволяет получать пространственные спектры изображений, затем выполнять традиционные для методов сжатия изображений усекающие действия в отношении наименее информативной части спектра и, наконец, устранять избыточность в оставшемся объеме информации. А широкое множество открытых квазиортогональных матриц, возможность «заказа» размерности таких матриц для выполнения преобразований в совокупности с возможностью применять процедуру перестановок строк и столбцов дают возможность использовать такой базис в системах защиты с закрытым ключом.

Заключение

Термин как слово, обозначающее определенное понятие, должен характеризоваться однозначностью. На практике одно и то же слово или словосочетание получает спектр значений, находя применение в разных областях знаний. Например, рассмотренный термин «маскирование» имеет не менее восьми специализированных значений, помимо словарного. Распирение понятия отдельных слов — процесс неизбежный, поскольку увели-

чивается количество направлений исследований, требующих новых определений, а толковые словари пополняются намного медленнее.

В отношении применения термина «маскирование» в области защиты цифровой визуальной информации были определены факторы, способствующие его закреплению в данном исследовательском и прикладном направлении «де-юре». Приведен также ряд отличий и ограничений, которые формально способны препятствовать признанию термина, хотя они и не носят категорического характера.

По мнению авторов настоящей работы, в пользу действительной годности термина «маскирование» в данном случае выступают следующие факторы:

- в языке отсутствует сколь-нибудь более близкая терминология с определением, способным указать точное место рассматриваемых процедур среди способов защиты информации;
- по крайней мере, одно из всех словарных определений термина напрямую обозначает цель процедур маскирования в применении к цифровой визуальной информации;
- термин успешно прошел многочисленную апробацию в результате применения в более чем 20 опубликованных работах в рецензируемых изданиях и представленных докладах на международных конференциях.

Литература

1. Ерош И. Л., Сергеев М. Б. Скоростное шифрование разнородных сообщений // Вопросы передачи и защиты информации: сб. ст. СПб.: СПбГУАП, 2006. С. 133–155.
2. Литвинов М. Ю. Алгоритмы маскирующих преобразований видеоинформации: автореф. дис. ... канд. техн. наук. — СПб.: ГУАП, 2009. — 23 с.
3. Беззатеев С. В., Литвинов М. Ю., Трояновский Б. К., Филатов Г. П. Выбор алгоритма преобразования, обеспечивающего изменение структуры изображений // Информационно-управляющие системы. 2006. № 6(25). С. 2–6.
4. Балонин Н. А., Сергеев М. Б. М-матрицы // Информационно-управляющие системы. 2011. № 1(50). С. 14–21.
5. Научно-технический энциклопедический словарь. <http://ucheba.ru/dictionary/word/1120538/> (дата обращения: 26.04.2014).
6. Словарь военных терминов / сост. А. М. Плехов, С. Г. Шапкин. — М.: Воениздат, 1988. — 336 с.
7. Химическая энциклопедия / под ред. И. Л. Кнунянца. — М.: Сов. энциклопедия, 1988. — 623 с.
8. Оксфордский толковый словарь по психологии: в 2 т. / под ред. А. Ребера; пер. с англ. Е. Ю. Чеботарева. — М.: Вече АСТ, 2003. Т. 1. — 592 с.
9. Большой психологический словарь / под ред. Б. Г. Мещерякова, акад. В. П. Зинченко. — М.: Прайм-ЕВРОЗНАК, 2003. — 672 с.
10. Сайт компании Oracle. <http://www.oracle.com/> (дата обращения: 04.05.2014).
11. SecurityLab.ru, IBM представила ПО для маскировки закрытых данных. <http://www.securitylab.ru/news/301841.php> (дата обращения: 03.05.2014).
12. Железняк В. К., Барков А. В. Экспериментальное исследование метода адаптивного маскирования видеосигнала от утечки по техническим каналам // Вестник Полоцкого государственного университета. Сер. С. Фундаментальные науки. 2014. № 4. С. 18–23.
13. kunegin.narod.ru, Маскиратор. <http://kunegin.narod.ru/ref8/shifr/mas.htm> (дата обращения: 03.08.2015).
14. Непрезкое маскирование. http://ru.wikipedia.org/wiki/%CD%E5%F0%E5%E7%EA%EE%E5_%EC%E0%F1%EA%E8%F0%EE%E2%E0%ED%E8%E5 (дата обращения: 11.05.2014).
15. Справка по Photoshop. Разд. «Маскирование слоев». <http://helpx.adobe.com/ru/photoshop/using/masking-layers.html> (дата обращения: 11.05.2014).
16. Маска, фр. masque — этимология. <http://www.proza.ru/2013/05/07/483> (дата обращения: 03.08.2015).
17. Ефремова Т. Ф. Современный словарь русского языка три в одном: орфографический, словообра-

- зовательный, морфемный. — М.: АСТ, 2010. — 699 с.
18. Валгина Н. С. Активные процессы в современном русском языке: учеб. пособие. — М.: Логос, 2001. — 304 с.
 19. Востриков А. А., Чернышев С. А. Об оценке устойчивости к искажениям изображений, маскированных М-матрицами // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 5. С. 99–103.
 20. Востриков А. А. О матрицах Адамара — Мерсенна и маскировании изображений // Информационные технологии. 2013. № 11. С. 37–39.
 21. Балонин Ю. Н., Востриков А. А. Матрицы Адамара — Мерсенна как базис ортогональных преобразований при маскировании видеоизображений // Приборостроение. 2014. № 1. С. 15–19.
 22. Востриков А. А., Мишура О. В., Сергеев А. М., Чернышев С. А. О выборе матриц для процедур маскирования и демаскирования изображений // Фундаментальные исследования. 2015. № 2 (ч. 24). С. 5335–5339.
 23. Балонин Н. А., Сергеев М. Б., Мироновский Л. А. Вычисление матриц Адамара — Мерсенна // Информационно-управляющие системы. 2012. № 5(60). С. 92–94.
 24. Балонин Н. А., Сергеев М. Б., Мироновский Л. А. Вычисление матриц Адамара — Ферма // Информационно-управляющие системы. 2012. № 6(61). С. 90–93.
 25. Балонин Н. А., Сергеев М. Б. О двух способах построения матриц Адамара — Эйлера // Информационно-управляющие системы. 2013. № 1(62). С. 7–10.
 26. Балонин Н. А., Балонин Ю. Н., Востриков А. А., Сергеев М. Б. Вычисление матриц Мерсенна — Уолша // Вестник компьютерных и информационных технологий. 2014. № 11. С. 51–55.
 27. Сергеев А. М. Обобщенные матрицы Мерсенна и гипотеза Балонина // Автоматика и вычислительная техника. 2014. № 4. С. 35–43.
 28. Балонин Н. А., Сергеев М. Б. К вопросу существования матриц Мерсенна и Адамара // Информационно-управляющие системы. 2013. № 5(66). С. 2–8.
 29. Балонин Н. А., Сергеев М. Б. Матрицы локального максимума детерминанта // Информационно-управляющие системы. 2014. № 1(68). С. 2–15.
 30. Балонин Ю. Н., Востриков А. А., Сергеев М. Б. О прикладных аспектах применения М-матриц // Информационно-управляющие системы. 2012. № 1(56). С. 92–93.
 31. Vostrikov A., Chernyshev S. Implementation of Novel Quasi-Orthogonal Matrices for Simultaneous Images Compression and Protection // Frontiers in Artificial Intelligence and Applications. 2014. Vol. 262, «Smart Digital Futures». P. 451–461. doi:10.3233/9781614994053451

UDC 004.052.2

doi:10.15217/issn1684-8853.2015.5.116

Masking of Digital Visual Data: the Term and Basic DefinitionsVostrikov A. A.^a, PhD, Tech., Associate Professor, vostricov@mail.ruSergeev M. B.^a, Dr. Sc., Tech., Professor, mbse@mail.ruLitvinov M. Yu.^b, PhD, Tech., Associate Professor^aDepartment of Computing Systems and Networks Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation^bBaumanMoscowStateTechnicalUniversity, 5/1, ul. Baumanskaya 2-ya, 105005, Moscow, Russian Federation

Purpose: At present, there are several terms to describe symmetric conversions of digital visual information in order to conceal it from unauthorized review. Significant differences in the implementation of conversions, different mathematical methods and respective demands on the efficiency make it necessary that the terms are in compliance with the procedures. Our goal is to take stock of the terms "masking" and "demasking", to consider alternative terminology for conversions performed when protecting digital images with a short relevance, and to fix the most appropriate option. **Results:** A comparative analysis of terminology close in meaning to the description of the contents of digital visual information converters showed that at the moment there are no more accurate terms than "masking" and "demasking", with a definition that can specify the exact place of the discussed procedures among the information protection methods. **Practical relevance:** It is shown that the term "masking" and the associated term "demasking" are closest to the conceptual description of digital video data protective conversion with a short relevance. Basic definitions are formulated for matrix conversions of digital visual information using quasi-orthogonal bases. Matrix equations are given for the implementation of "masking" and "unmasking".

Keywords — Visual Data Protection, Image Protection, Masking, Unmasking, Quasi-Orthogonal Bases, Quasi-Orthogonal Matrix, Cryptographic Primitives.

References

1. Erosh I. L., Sergeev M. B. High-Speed Encryption Disparate Messages. *Voprosy peredachi i zashchity informatsii* [Transmission and Protection of Information]. Saint-Petersburg, GUAP Publ., 2006, pp. 133–155 (In Russian).
2. Litvinov M. Y. *Algoritmy maskiruiushchikh preobrazovaniy videoinformatsii*. Dis. kand. tehn. nauk [Algorithms of Video Masking Transforms. PhD tech. sci. diss.]. Saint-Petersburg, GUAP Publ., 2006. 23 p. (In Russian).

3. Bezzateev S. V., Litvinov M. Y., Troyanovskii B. K., Filatov G. P. The Choice of the Transformation Algorithm that Ensures a Structural Change of Videoinformation. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2006, no. 6(25), pp. 2–6 (In Russian).
4. Balonin N. A., Sergeev M. B. M-Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2011, no. 1(50), pp. 14–21 (In Russian).
5. *Nauchno-tekhnicheskii entsiklopedicheskii slovar'* [Scientific and Technical Encyclopedic Dictionary]. Available at: <http://ucheba.ru/dictionary/word/1120538/> (accessed 26 April 2014).
6. Plehov A. M., Shapkin S. G. *Slovar' voennykh terminov* [Dictionary of Military Terms]. Moscow, Voenizdat Publ., 1988. 336 p. (In Russian).
7. Knujanc I. L. *Khimicheskaja entsiklopedija* [Chemical Encyclopedia]. Moscow, Sovetskaja entsiklopedija Publ., 1988. 336 p. (In Russian).
8. Reber A. *Oksfordskii tolkovyi slovar' po psikhologii* [The Oxford Dictionary of Psychology]. Moscow, Veche AST Publ., 2003. Vol. 2. 592 p. (In Russian).
9. Meshherjakova B. G., Zinchenko V. P. *Bol'shoi psikhologicheskii slovar'* [The Big Psychological Dictionary]. Moscow, Prain-EVROZNAK Publ., 2003. 672 p. (In Russian).
10. *Sait kompanii Oracle* [Saite of Company Oracle]. Available at: <http://www.oracle.com/> (accessed 04 May 2014).
11. *IBM predstavila PO dlia maskirovki zakrytykh dannykh* [IBM Introduced the Software to Mask Sensitive Data]. Available at: <http://www.securitylab.ru/news/301841.php> (accessed 03 May 2014).
12. Zhelezniak V. K., Barkov A. V. Experimental Study of Adaptive Video Masking Method for Protection Against Leakage Through Technical Channels. *Vestnik Polotskogo gosudarstvennogo universiteta. Ser. C. Fundamental'nye nauki*, 2014, no. 4, pp. 18–23 (In Russian).
13. *Maskirator* [Scrambler]. Available at: <http://kunegin.narod.ru/ref8/shifr/mas.htm> (accessed 03 August 2015).
14. *Nerezko maskirovanie* [Unsharp Masking]. Available at: http://ru.wikipedia.org/wiki/%CD%E5%F0%E5%E7%EA%EE%E5_%EC%E0%F1%EA%E8%F0%EE%E2%E0%ED%E8%E5 (accessed 11 May 2015).
15. *Spravka po Photoshop. Razdel "Maskirovanie sloev"* [Help Photoshop. Section "Masking Layers"]. Available at: <http://helpx.adobe.com/ru/photoshop/using/masking-layers.html> (accessed 11 May 2015).
16. *Maska, fr. masque — etimologija* [Mask fr. masque — Etymology]. Available at: <http://www.proza.ru/2013/05/07/483> (accessed 03 August 2015).
17. Efremova T. F. *Sovremennyi slovar' russkogo iazyka tri v odnom: orfograficheskii, slovoobrazovatel'nyi, morfemnyi* [Modern Dictionary of the Russian Language in One of Three: Spelling, Word-Formation, Morphemic]. Moscow, AST Publ., 2010. 699 p. (In Russian).
18. Valgina N. S. *Aktivnye protsessy v sovremennom russkom iazyke* [Active Processes in Modern Russian]. Moscow, Logos Publ., 2001. 304 p. (In Russian).
19. Vostrikov A. A., Chernyshev S. A. On Distortion Assessment of Images Masking with M-Matrices. *Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologii, mekhaniki i optiki*, 2013, no. 5, pp. 99–103 (In Russian).
20. Vostrikov A. A. On Hadamard–Mersenne Matrices and Image Masking. *Informatsionnye tekhnologii*, 2013, no. 11, pp. 37–39 (In Russian).
21. Balonin Y. N., Vostrikov A. A. Hadamard–Mersenne Matrices as a Basis of Orthogonal Transformation for Video Masking Encoding. *Priboroostroenie*, 2014, pp. 15–19 (In Russian).
22. Vostrikov A. A., Mishura O. V., Sergeev A. M., Chernyshev S. A. The Choice of Matrices for Images Masking and Demasking Procedures. *Fundamental'nye issledovaniia*, 2015, no. 2 (part 24), pp. 5335–5339 (In Russian).
23. Balonin N. A., Sergeev M. B., Mironovsky L. A. Calculation of Hadamard–Mersenne Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 5(60), pp. 92–94 (In Russian).
24. Balonin N. A., Sergeev M. B., Mironovsky L. A. Calculation of Hadamard–Fermat Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 6(61), pp. 90–93 (In Russian).
25. Balonin N. A., Sergeev M. B. Two Ways to Construct Hadamard–Euler Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 1(62), pp. 7–10 (In Russian).
26. Balonin N. A., Balonin Yu. N., Vostrikov A. A., Sergeev M. B. Computation of Mersenne–Walsh Matrices. *Vestnik komp'iuternykh i informatsionnykh tekhnologii*, 2014, no. 11, pp. 51–55 (In Russian).
27. Sergeev A. M. The Generalized Matrices of Mersenne and Balonin's Conjecture. *Avtomatika i vychislitel'naja tekhnika*, 2014, no. 4, pp. 35–43 (In Russian).
28. Balonin N. A., Sergeev M. B. On the Issue of Existence of Hadamard and Mersenne Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2013, no. 5(66), pp. 2–8 (In Russian).
29. Balonin N. A., Sergeev M. B. Local Maximum Determinant Matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2014, no. 1(68), pp. 2–15 (In Russian).
30. Balonin Yu. N., Vostrikov A. A., Sergeev M. B. Applied Aspects of M-Matrix Use. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2012, no. 1(56), pp. 92–93 (In Russian).
31. Vostrikov A., Chernyshev S. Implementation of Novel Quasi-Orthogonal Matrices for Simultaneous Images Compression and Protection. *Frontiers in Artificial Intelligence and Applications*, 2014, vol. 262, "Smart Digital Futures", pp. 451–461. doi:10.3233/9781614994053451

ПАРЕТО-ОПТИМАЛЬНОСТЬ В СТАТИЧЕСКОЙ КОНКУРЕНТНОЙ МОДЕЛИ ПРИНЯТИЯ РЕШЕНИЙ

К. В. Григорьева^а, канд. физ.-мат. наук, доцент

^аСанкт-Петербургский государственный университет, Санкт-Петербург, РФ

Постановка проблемы: в ряде прикладных задач, таких как задачи прогнозирования, выбора, назначения и распределения, диагностики и многоагентного управления и др., иногда возникает проблема построения оптимального взаимодействия между агентами. **Цель:** построение нового алгоритма решения для теоретико-игровой модели многоагентного взаимодействия конкурентного типа с использованием парето-оптимальности и компромиссного множества, который позволит обрабатывать данные (проводить анализ данных) большого количества участников в каждом проекте с помощью построения несложного программного обеспечения. **Результаты:** построен алгоритм решения статической конкурентной модели принятия решений, заключающийся в поиске парето-оптимального решения в бескоалиционных играх и компромиссного проекта. Статическая конкурентная модель принятия решений формализуется в виде множества различных между собой бескоалиционных игр, каждая из которых задана для некоторого проекта. Для каждого проекта в качестве стратегий игроков выступают положительное и отрицательное решение по соответствующему проекту. Доходы игроков определяются как значения функций выигрыша на множестве ситуаций, образованных принятыми решениями игроков по соответствующим проектам. Требуется решить каждую бескоалиционную игру, а затем из множества полученных решений выделить компромиссное с помощью алгоритма нахождения компромиссного решения в целях выделения приоритетного проекта (одного или нескольких). Доказано существование решения статической конкурентной модели принятия решений, приведен численный пример ее реализации. **Практическая значимость:** предложенный алгоритм может быть рекомендован к использованию для экспертов как инструмент для уточнения или подтверждения оптимальности предполагаемого решения по участию в том или ином проекте.

Ключевые слова — бескоалиционная игра, парето-оптимальное решение, арбитражная схема Нэша, компромиссное решение, многоагентные системы.

Введение

Задача, рассматриваемая в этой статье, относится к теории принятия решений [1], и если более конкретно, то к серии задач наилучшего выбора [2–5]. Проблема решения задач наилучшего выбора в различных постановках остается актуальной в силу повседневности и многообразия применения этих задач в различных областях жизнедеятельности. Разработка моделей и аппарата для решения задач этого класса в рамках теории управления и теории принятия решений является востребованной и в настоящее время. Суть исследуемой задачи заключается в следующем. Пусть заданы множество агентов N и множество проектов M . Агенты принимают решение: участвовать или не участвовать в проекте. Участие в проекте сопряжено с доходами или расходами, которые агенты хотят, соответственно, максимизировать или минимизировать. Необходимо найти проект (один или несколько), участие в котором принесло бы агентам наилучший в некотором смысле доход. Очевидно, что это — оптимизационная задача, и ее решение может быть найдено путем построения теоретико-игровой модели.

Для решения поставленной задачи построим модель, состоящую из m бескоалиционных игр n лиц G_j , $j \in M$, где $n = |N|$ — количество игроков, а $m = |M|$ — количество проектов. В каждой

игре G_j игрок $i \in N$ имеет две стратегии: принять или отклонить проект $j \in M$. Каждый игрок делает свой выбор, и образуется ситуация игры. Выигрыш каждого игрока определяется на множестве ситуаций в этой игре.

Однако задача формулируется так, что необходимо выбрать подмножество из проектов, приносящих оптимальные доходы для игроков. Для этого решим поочередно каждую из m бескоалиционных игр, получим набор оптимальных ситуаций и набор векторов выигрышей в этих ситуациях в соответствующих бескоалиционных играх. Применим к полученному набору векторов выигрышей в оптимальных ситуациях алгоритм нахождения компромиссного множества [6] и получим один или несколько компромиссных проектов.

Для решения бескоалиционных игр в этой статье предлагается использовать парето-оптимальное решение и арбитражную схему Нэша.

До сих пор такие модели решались путем построения одной большой игры, в которой каждая ситуация включает в себя мнение всех игроков по каждому проекту, что является нерациональным. Парето-оптимальность же ко всему прочему еще и обеспечивает наличие решения в бескоалиционных играх с конечным множеством ситуаций всегда.

Ранее в работах [7–11] автором рассматривались подобные статические модели принятия

решений с коалиционными разбиениями. Эти модели различаются по следующим критериям: в работе [7] для всех проектов задана одна бескоалиционная игра, но для каждого проекта с этой игрой ассоциировано свое коалиционное разбиение; в работе [8] для каждого проекта задана своя бескоалиционная игра и ассоциированное с этой игрой коалиционное разбиение; статьи [9] и [10] являются обобщением работы [7] на многокритериальный случай задания выигрышей игроков, и проблема выбора игроками приемлемого в некотором смысле проекта решается путем сведения многокритериальной коалиционной модели принятия решений с постоянной матрицей выигрышей в бескоалиционной игре к однокритериальной статической модели принятия решений с единственной бескоалиционной игрой и фиксированными коалиционными разбиениями с помощью метода взвешенных коэффициентов многокритериальной оптимизации в первом случае и с помощью минимаксного метода многокритериальной оптимизации во втором.

В данной работе предлагается рассмотреть случай однокритериальной статической модели принятия решений без коалиционных разбиений. Будем называть ее *статической конкурентной моделью принятия решений*.

Приведем примеры возможного применения этой модели.

1. Предположим, что мы имеем несколько инновационных проектов и набор инвесторов. Зафиксируем инновационный проект. Каждый инвестор решает: принимать участие в этом проекте или нет. Доходы каждого инвестора определяются на множестве ситуаций, образованных в соответствии с решениями всех инвесторов. Мы должны выбрать инновационный проект (один или несколько), который будет в каком-то смысле более привлекательным для инвесторов.

2. Пусть в законодательном органе предлагается несколько вариантов какого-либо законопроекта, который обсуждается законодателями. В этом примере агентами являются законодатели, проектами — варианты законопроекта. Стратегиями законодателей являются одобрение версии законопроекта или принятие решения голосовать против предложенного варианта. Доходы каждого законодателя определяются на множестве ситуаций, образующихся из решений всех законодателей при фиксированном варианте законопроекта. Необходимо выбрать одну или несколько версий законопроекта, которые наиболее справедливы и учитывают все аспекты негативных последствий вступления закона в силу.

3. Существует несколько проектов переселения жильцов из аварийного дома. Жильцам предоставляется выбор нового места жительства

из нескольких предлагаемых вариантов. В этом примере агентами являются жильцы, проектами — конкретные объекты переселения, стратегиями агентов — согласие или отказ переселиться в соответствующее здание. Для того чтобы избежать конфликта с жильцами, эксперт рекомендует застройщику место или параметры здания по результатам применения предложенной модели.

В статье приведен численный пример реализации предложенной модели с участием трех агентов.

Постановка задачи и необходимые теоретические сведения для ее решения

Пусть имеются множество $N = \{1, \dots, n\}$ игроков и множество $P = \{1, \dots, m\}$ проектов, по которым требуется принятие положительного или отрицательного решения каждым из игроков.

Введем обозначения:

$X_i = \{0; 1\}$ — множество чистых стратегий x_i игрока $i \in N$; стратегия x_i соответственно может принимать значения: $x_i = 0$ — отрицательное решение игрока i по некоторому проекту, $x_i = 1$ — положительное решение;

$x = (x_1, \dots, x_n)$ — ситуация игры между игроками в чистых стратегиях;

$X = \prod_{i=1, n} X_i$ — множество ситуаций игры.

На множестве ситуаций в чистых стратегиях для каждого проекта $j \in M$ и для каждого игрока $i \in N$ определена функция выигрыша $K_i^j: X \rightarrow R^1$. Вектор выигрышей игроков будем обозначать через $K^j(x) = \{K_i^j(x)\}_{i=1}^n$.

Таким образом, имеем m бескоалиционных игр n лиц $G_j(x)$, $j \in M$:

$$G_j(x) = \left\langle N, \{X_i\}_{i=1, n}, \{K_i^j(x)\}_{i=1, n}, x \in X \right\rangle. \quad (1)$$

Под *решением бескоалиционной игры* $G_j(x)$, $j \in M$, будем понимать подмножество парето-оптимальных ситуаций, выделенных из множества всех парето-оптимальных ситуаций с помощью арбитражной схемы Нэша.

Напомним определения парето-оптимального множества (PO) [12–14] и арбитражного решения Нэша (NB) [14–16].

Определение 1. Ситуация $\bar{x}$ в бескоалиционной игре называется *оптимальной по Парето*, если не существует никакой другой ситуации x' такой, что:

- 1) $K_i(x') \geq K_i(\bar{x}) \forall i \in N$;
- 2) хотя бы для одного $i_0 \in N$ $K_{i_0}(x') > K_{i_0}(\bar{x})$.

Обозначим множество парето-оптимальных ситуаций через $\bar{X}$: $\bar{X} \subset X$.

Определение 2. Множество $\bar{X} \subset X$ называется *множеством допустимых результатов*.

Определение 3. Арбитражным решением Нэша $\bar{x}$ называется решение следующей оптимизационной задачи [16]:

$$\max_{x \in \bar{X}} \prod_{i=1}^n [K_i(x) - v_i] = \prod_{i=1}^n [K_i(\bar{x}) - v_i]$$

при ограничениях

$$K_i(x) \geq v_i \quad \forall i \in N.$$

Здесь $v = (v_1, \dots, v_n)$ — точка «статус-кво», обозначающая исход бескоалиционной игры, который игроки себе гарантируют; этой точке соответствует ситуация x' :

$$v_i = v_i(x') \quad \forall i \in N.$$

В качестве v_i будем использовать максимальный гарантированный выигрыш i -го игрока, определяемый характеристической функцией в кооперативной игре n лиц [12]:

$$v_i = \max_{x_i \in X_i} \min_{x_{N \setminus \{i\}} \in X_{N \setminus \{i\}}} K_i(x_i, x_{N \setminus \{i\}}), \quad (2)$$

где $x_{N \setminus \{i\}} = (x_i, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ — стратегия коалиции $N \setminus \{i\}$.

Пусть для каждого проекта $j \in M$ найдено решение $\bar{x}$ бескоалиционной игры $G_j(x)$.

Определение 4. Компромиссным множеством называется множество

$$C_K(P) = \arg \min_j \max_i \left\{ \max_j K_i^j(\bar{x}) - K_i^j(\bar{x}) \right\},$$

где $K = \{K_i^j(\bar{x})\}_{i=1, \dots, n, j=1, \dots, m}$ — матрица выигрышей игроков в NB $\bar{x}$ [6].

Таким образом, компромиссным решением является «наиболее удобный» проект, мера «удобства» которого определяется минимальностью из всех максимальных отклонений выигрыша по всем игрокам для каждого отдельного проекта; при этом под отклонением выигрыша понимается значение недобора выигрыша игрока до его максимального возможного выигрыша.

Определение 5. Полным компромиссным множеством называется множество

$$Compr_K(P) = C_K^l(P),$$

где $C_K^l(P)$ определяется следующим образом:

$$C_K^1(P) = C_K(P);$$

$$C_K^q(P) = \arg \min_{j \in C_K^{q-1}(P)} \max_i \left\{ \max_{j \in C_K^{q-1}(P)} K_i^j(\bar{x}) - K_i^j(\bar{x}) \right\},$$

$$q = \overline{2, l+1}; \quad C_K^l(P) = C_K^{l+1}(P).$$

Проект $\bar{j} \in Compr_K(P)$ является компромиссным решением для всех игроков.

Определение 6. Статической конкурентной моделью принятия решений Γ будем называть множество из m бескоалиционных игр $G_j(x)$, $j \in M$, определенных выражением (1):

$$\Gamma = \{G_j(x)\}_{j=1}^m.$$

Определение 7. Под решением статической конкурентной модели принятия решений будем понимать пару $(\bar{x}, \bar{j})$, где $\bar{x}$ — NB-решение в бескоалиционной игре $G_{\bar{j}}(x)$ с компромиссным проектом $\bar{j}$.

Определение 8. Вектор выигрышей $K^{\bar{j}}(\bar{x})$ игроков в ситуации NB будем называть компромиссным вектором выигрышей игроков.

Алгоритм принятия решения статической конкурентной модели и численный пример его реализации

Имеем алгоритм решения модели:

1. Зафиксируем $j \in M$.
2. Найдём множество РО-решений $\bar{X}$ в бескоалиционной игре $G_j(x)$.
3. Найдём множество NB-решений $\bar{x} \in \bar{X}$ в бескоалиционной игре $G_j(x)$.
4. Повторим итерации 1–3 для всех остальных $j \in M$.
5. Найдём полное компромиссное множество проектов $\bar{j} \in Compr_K(P)$.

Переформулируем теорему из работы [17, с. 39] в наших обозначениях.

Теорема. В случае конечного множества возможных решений X существует хотя бы одно парето-оптимальное решение, т. е. $\bar{X} \neq \emptyset$.

Утверждение. Статическая конкурентная модель принятия решений Γ всегда имеет по крайней мере одно решение, полученное с помощью алгоритма решения статической конкурентной модели принятия решений.

Доказательство: Действительно, согласно теореме, в бескоалиционной игре с конечным множеством ситуаций X хотя бы одна РО-ситуация существует всегда. Арбитражное решение Нэша на непустом множестве РО-решений также существует всегда, поскольку всегда найдется максимум от дискретной функции выигрыша (определение 3). Наконец, полное компромиссное множество всегда не пусто, поскольку всегда существуют минимум и максимум от дискретной функции выигрыша (определение 5).

Пример

Рассмотрим множество $P = \{j\}_{j \in 1:2}$ и множество из трех игроков $N = \{i\}_{i \in 1:3}$, у каждого из которых в бескоалиционной игре $G_j(x)$ множество стратегий содержит два элемента: $x_i = 1$ — «да»,

$x_i = 0$ — «нет», $i \in 1 : 3$. Функции выигрыша игроков определяются в играх $G_j(x)$ посредством табл. 1.

1. Решим бескоалиционную игру $G_1(x)$. В чистых стратегиях равновесия по Нэшу не существует.

Вычислим с помощью формулы (2) гарантированные выигрыши игроков:

$$v_1 = \max_{x_1 \in X_1} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_1} \in \tilde{X}_{N \setminus I_1}} \{4; 1; 3; 5\}; \\ \min_{\tilde{x}_{N \setminus I_1} \in \tilde{X}_{N \setminus I_1}} \{5; 1; 0; 0\} \end{array} \right\} = \max_{x_1 \in X_1} \{1, 0\} = 1;$$

$$v_2 = \max_{x_2 \in X_2} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_2} \in \tilde{X}_{N \setminus I_2}} \{2; 2; 3; 2\}; \\ \min_{\tilde{x}_{N \setminus I_2} \in \tilde{X}_{N \setminus I_2}} \{1; 1; 4; 4\} \end{array} \right\} = \max_{x_2 \in X_2} \{2, 1\} = 2;$$

$$v_3 = \max_{x_3 \in X_3} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_3} \in \tilde{X}_{N \setminus I_3}} \{1; 5; 1; 3\}; \\ \min_{\tilde{x}_{N \setminus I_3} \in \tilde{X}_{N \setminus I_3}} \{2; 3; 2; 2\} \end{array} \right\} = \max_{x_3 \in X_3} \{1, 2\} = 2.$$

Найдем оптимальные стратегии согласно арбитражной схеме Нэша (табл. 2), где «—» означает, что стратегии не оптимальны по Парето, а «+» — оптимальны по Парето. Тогда оптимальными будем считать ситуации (1,1,0) и (0,1,0), которые дают одинаковый выигрыш (1,2,2) в обеих ситуациях.

2. Решим бескоалиционную игру $G_2(x)$ (см. табл. 1). В чистых стратегиях равновесия по Нэшу не существует.

Вычислим гарантированные выигрыши игроков:

$$v_1 = \max_{x_1 \in X_1} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_1} \in \tilde{X}_{N \setminus I_1}} \{4; 1; 2; 5\}; \\ \min_{\tilde{x}_{N \setminus I_1} \in \tilde{X}_{N \setminus I_1}} \{5; 1; 0; 0\} \end{array} \right\} = \max_{x_1 \in X_1} \{1, 0\} = 1;$$

$$v_2 = \max_{x_2 \in X_2} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_2} \in \tilde{X}_{N \setminus I_2}} \{2; 2; 3; 2\}; \\ \min_{\tilde{x}_{N \setminus I_2} \in \tilde{X}_{N \setminus I_2}} \{1; 1; 3; 4\} \end{array} \right\} = \max_{x_2 \in X_2} \{2, 1\} = 2;$$

$$v_3 = \max_{x_3 \in X_3} \left\{ \begin{array}{l} \min_{\tilde{x}_{N \setminus I_3} \in \tilde{X}_{N \setminus I_3}} \{1; 5; 1; 3\}; \\ \min_{\tilde{x}_{N \setminus I_3} \in \tilde{X}_{N \setminus I_3}} \{2; 3; 3; 2\} \end{array} \right\} = \max_{x_3 \in X_3} \{1, 2\} = 2.$$

Найдем оптимальные стратегии согласно арбитражной схеме Нэша (табл. 3), где «—» означает, что стратегии не оптимальны по Парето, а «+» — оптимальны по Парето. Тогда оптимальной будем считать ситуацию (0,1,0), которая дает выигрыш (1,2,3).

■ Таблица 1

Стратегия игрока			Выигрыш в игре $G_1(x)$ игрока			Выигрыш в игре $G_2(x)$ игрока		
1	2	3	1	2	3	1	2	3
1	1	1	4	2	1	4	2	1
1	1	0	1	2	2	1	2	2
1	0	1	3	1	5	2	1	5
1	0	0	5	1	3	5	1	3
0	1	1	5	3	1	5	3	1
0	1	0	1	2	2	1	2	3
0	0	1	0	4	3	0	3	3
0	0	0	0	4	2	0	4	2

■ Таблица 2

Стратегия игрока			Выигрыш игрока			Арбитражная схема Нэша	Оптимальность по Парето
1	2	3	1	2	3		
1	1	1	4	2	1	$(4-1)(2-2)(1-2) = 0$	—
1	1	0	1	2	2	$(1-1)(2-2)(2-2) = 0$	+
1	0	1	3	1	5	$(3-1)(1-2)(5-2) < 0$	—
1	0	0	5	1	3	$(5-1)(1-2)(3-2) < 0$	—
0	1	1	5	3	1	$(5-1)(3-2)(1-2) < 0$	—
0	1	0	1	2	2	$(1-1)(2-2)(2-2) = 0$	+
0	0	1	0	4	3	$(0-1)(4-2)(3-2) < 0$	—
0	0	0	0	4	2	$(0-1)(4-2)(2-2) = 0$	—

■ Таблица 3

Стратегия игрока			Выигрыш игрока			Арбитражная схема Нэша	Оптимальность по Парето
1	2	3	1	2	3		
1	1	1	4	2	1	$(4-1)(2-2)(1-2) = 0$	—
1	1	0	1	2	2	$(1-1)(2-2)(2-2) = 0$	—
1	0	1	2	1	5	$(2-1)(1-2)(5-2) < 0$	+
1	0	0	5	1	3	$(5-1)(1-2)(3-2) < 0$	+
0	1	1	5	3	1	$(5-1)(3-2)(1-2) < 0$	+
0	1	0	1	2	3	$(1-1)(2-2)(3-2) = 0$	+
0	0	1	0	3	3	$(0-1)(3-2)(3-2) < 0$	+
0	0	0	0	4	2	$(0-1)(4-2)(2-2) = 0$	+

Применяя алгоритм нахождения компромиссного множества, получаем множество компромиссных проектов (табл. 4, 5).

Следовательно, компромиссным является 2-й проект. При этом в ситуации (0,1,0) первый и третий агенты дают отрицательное решение по соответствующему проекту. Иными словами,

■ Таблица 4

$\bar{x}$	K_1^j	K_2^j	K_3^j
(1,1,0); (0,1,0)	1	2	2
(0,1,0)	1	2	3
M_i	1	2	3

■ Таблица 5

$\bar{x}$	(1,1,0); (0,1,0)	(0,1,0)
$M_1 - K_1^j$	0	0
$M_2 - K_2^j$	0	0
$M_3 - K_3^j$	1	0
$\max_i \{M_i - K_i^j\}$	1	0

если первый и третий агенты дают отрицательное решение по соответствующему проекту, а второй — нет, то их доход окажется оптимальным.

Литература

1. Kolbin V. V. Decision Making and Programming. — World Scientific Publishing Co, 2003. — 756 p.
2. Yuh-Wen Chen. A Group Game of Multiple Attribute Decision Making // Asia-Pacific Journal of Operational Research. 2007. N 24(5). P. 631–645.
3. Ayeley P. Tchangani. Evaluation Model for Multi Attributes — Multi Agents Decision Making. Satisficing Game Approach // International Journal of Information Technology and Decision Making. 2009. N 8(1). P. 73–91.
4. Sajjad Zahir, Ruhul A. Sarker. Planning Relocation of People for Developing Surface Mines in Densely Populated Area. Optimization of Multiple Objectives // Asia-Pacific Journal of Operational Research. 2011. N 28(5). P. 563–583.
5. Трахтенгерц Э. А. Взаимодействие агентов в много-агентных системах // Автоматика и Телемеханика. 1998. Вып. 8. С. 3–52.
6. Григорьева К. В., Малафеев О. А. Методы решения динамической многокритериальной задачи почтальона // Вестник гражданских инженеров. 2011. № 4. С. 156–161.
7. Григорьева К. В. Статическая модель принятия решений с единственной бескоалиционной игрой и фиксированными коалиционными разбиениями // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Сер. 1. Естественные и технические науки. 2013. № 1. С. 29–36.
8. Григорьева К. В. Статическая модель принятия решений с различными бескоалиционными играми и фиксированными коалиционными разбиениями // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Сер. 1. Естественные и технические науки. 2013. № 3. С. 24–30.
9. Григорьева К. В. Многокритериальная коалиционная модель принятия решений с постоянной матрицей выигрышей в бескоалиционной игре // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Сер. 1. Естественные и технические науки. 2013. № 3. С. 32–36.
10. Григорьева К. В. Статическая модель принятия решений с постоянной матрицей выигрышей в бескоалиционной игре и фиксированными коалиционными разбиениями с использованием минимаксного метода многокритериальной оптимизации // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Сер. 1. Естественные и технические науки. 2014. № 1. С. 43–47.
11. Григорьева К. В. Коалиционная модель принятия решений на множестве проектов // Материалы конф. «Управление в технических, эргатических, организационных и сетевых системах», Санкт-Петербург, 9–11 октября 2012 г. СПб., 2012. С. 952–954.
12. Зенкевич Н. А., Петросян Л. А., Янг Д. В. К. Динамические игры и их приложения в менеджменте. — СПб.: Высшая школа менеджмента, 2009. — 415 с.
13. Григорьева К. В. Бескоалиционные игры в нормальной форме. — СПб.: СПбГАСУ, 2007. — 78 с.
14. Григорьева К. В. Конфликтно-динамические системы. Ч. 1: Статические и стохастические игры. — СПб.: ФГБОУ ВПО «СПГУТД», 2012. — 163 с.

Заключение

В статье предложена статическая конкурентная модель принятия решений, формализованная как множество различных между собой бескоалиционных игр, каждая из которых задана для некоторого проекта. Построен новый алгоритм решения статической конкурентной модели принятия решений, заключающийся в поиске парето-оптимального решения в бескоалиционных играх и компромиссного проекта; доказано существование решения статической конкурентной модели принятия решений; приведен численный пример ее реализации. В отличие от уже имеющихся моделей, решение задачи основано на оптимизационном подходе и рассчитано на большое количество участников. Для предложенной модели нетрудно создать программное обеспечение. Также предложенный в этой работе алгоритм может быть рекомендован к использованию для экспертов как инструмент для уточнения или подтверждения оптимальности предполагаемого решения.

15. Nash J. F. The Bargaining Problem // *Econometrica*. 1950. N 18(2). P. 155–162.
16. Зенкевич Н. А., Петросян Л. А. Проблема временной состоятельности кооперативных решений // *Научные доклады*. 2006. № 8(R). [http://www.gsom.spbu.ru/files/upload/niim/publishing/](http://www.gsom.spbu.ru/files/upload/niim/publishing/papers/Zenkevich_Petrosyan.pdf)

[papers/Zenkevich_Petrosyan.pdf](http://www.gsom.spbu.ru/files/upload/niim/publishing/papers/Zenkevich_Petrosyan.pdf) (дата обращения: 27.07.2015).

17. Ногин В. Д. Принятие решений в многокритериальной среде: количественный подход. — М.: Физматлит, 2004. — 176 с.

UDC 519.83; 518.9

doi:10.15217/issn1684-8853.2015.5.124

Pareto Optimality in a Static Competitive Decision-Making Model

Grigorieva X. V.^a, PhD, Phys.-Math., Associate Professor, k.grigorieva@spbu.ru

^aSaint-Petersburg State University, 35, Universitetskii Pr., 198504, Petergof, Saint-Petersburg, Russian Federation

Purpose: Some applied problems, such as forecasting, selecting, assignment, allocation, diagnostics or multi-agent management, often require providing optimal interaction between agents. The purpose of this research is constructing an algorithm and simple software for solving a game-theoretic model of multi-agent interaction of competitive type, using Pareto optimum and compromise solution, which would process data from a large number of participants in each project. **Results:** An algorithm is built for solving a static competitive decision-making model, which searches for Pareto-optimal solutions in non-cooperative games and for a compromise project. This model is formalized as a family of different non-cooperative games. Each game is defined for some project and requires a adoption of a positive or negative decision by every player. The players' incomes are defined as values of payoff functions on the set of n-tuples from their decisions for the relevant projects. We have to solve each non-cooperative game and then, from a set of solutions, to choose a compromise one with the help of an algorithm of finding a compromise solution in order to emphasize the priority project (one or more). The existence of static competitive decision-making model solution is proved and a numerical example is given. **Practical relevance:** The proposed algorithm can be recommended as a tool to clarify or confirm the optimal solution about the alleged participation in a certain project.

Keywords — Non-cooperative Game, Pareto Optimum, Nash Arbitration Solution, Compromise Solution, Multiagent Systems.

References

1. Kolbin V. V. *Decision making and Programming*. World Scientific Publishing Co, 2003. 756 p.
2. Yuh-Wen Chen. A Group Game of Multiple Attribute Decision Making. *Asia-Pacific Journal of Operational Research*, 2007, no. 24(5), pp. 631–645.
3. Ayeley P. Tchangani. Evaluation Model for Multi Attributes — Multi Agents Decision Making: Satisficing Game Approach. *International Journal of Information Technology and Decision Making*, 2009, no. 8(1), pp. 73–91.
4. Sajjad Zahir, Ruhul A. Sarker. Planning Relocation of People for Developing Surface Mines in Densely Populated Area. Optimization of Multiple Objectives. *Asia-Pacific Journal of Operational Research*, 2011, no. 28(5), pp. 563–583.
5. Trakhtengerts E. A. Interaction of Agents in Multi-agent Systems. *Avtomatika i Telemekhanika*, 1998, iss. 8, pp. 3–52 (In Russian).
6. Grigor'eva X. V., Malafeyev O. A. Methods for Solving the Dynamic Multicriteria Postman's Problem. *Vestnik grazhdanskikh inzhenerov*, 2011, no. 4, pp. 156–161 (In Russian).
7. Grigor'eva X. V. Static Model of Decision-making with a Single Noncooperative Game and Fixed Coalitional Partitions. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizaina. Ser. 1. Estestvennye i tekhnicheskie nauki*, 2013, no. 1, pp. 29–36 (In Russian).
8. Grigor'eva X. V. Static Model of Decision-making with a Different Noncooperative Games and Fixed Coalitional Partitions. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizaina. Ser. 1. Estestvennye i tekhnicheskie nauki*, 2013, no. 3, pp. 24–31 (In Russian).
9. Grigor'eva X. V. Multi-criteria Coalitional Model of Decision-making with a Constant Payoff Matrix in a Noncooperative Game. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizaina. Ser. 1. Estestvennye i tekhnicheskie nauki*, 2013, no. 3, pp. 32–36 (In Russian).
10. Grigor'eva X. V. Static Model of Decision-making with a Constant Payoff Matrix in a Noncooperative Game and Fixed Coalitional Partitions Using Minimax Method of Multi-criteria Optimization. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizaina. Ser. 1. Estestvennye i tekhnicheskie nauki*, 2014, no. 1, pp. 43–47 (In Russian).
11. Grigor'eva X. V. Coalitional Model of Decision-Making on a Set of Projects. *Materialy konferentsii "Upravlenie v tekhnicheskikh, ergaticheskikh, organizatsionnykh i setevykh sistemakh"* [Proc. of the Conference "Control in Technical, Ergatic, Organizational and Networking Systems"]. Saint-Petersburg, 2012, pp. 952–954 (In Russian).
12. Zenkevich N. A., Petrosjan L. A., Yeung D. W. K. *Dinamicheskie igry i ikh prilozheniia v menedzhmente* [Dynamic Games and Their Application in Management]. Saint-Petersburg, Vysshaia shkola menedzhmenta Publ., 2009. 415 p. (In Russian).
13. Grigor'eva X. V. *Beskoalitsionnye igry v normal'noi forme* [Noncooperative Games in Normal Form]. Saint-Petersburg, SPbGASU Publ., 2007. 78 p. (In Russian).
14. Grigor'eva X. V. *Konfliktno-dinamicheskie sistemy. Chast' 1. Staticheskie i stokhasticheskie igry* [Conflict Dynamic System. Part 1. Static and Stochastic Games]. Saint-Petersburg, SPbGASU Publ., 2012. 163 p. (In Russian).
15. Nash J. F. The Bargaining Problem. *Econometrica*, 1950, no. 18(2), pp. 155–162.
16. Zenkevich N. A., Petrosjan L. A. Time-consistency of Cooperative Solutions. *Nauchnye doklady*, 2006, no. 8(R). Available at: http://www.gsom.spbu.ru/files/upload/niim/publishing/papers/Zenkevich_Petrosyan.pdf (accessed 27 July 2015).
17. Nokin V. D. *Priniatie reshenii v mnogokriterial'noi srede: kolichestvennyi podkhod* [Decision-making in Multicriteria Environment: a Quantitative Approach]. Moscow, Fizmatlit Publ., 2004. 176 p. (In Russian).

**АРДАШОВ
Август
Анатолевич**



Старший научный сотрудник Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.

В 1961 году окончил Ростовское высшее артиллерийское инженерное училище им. Главного маршала артиллерии М. И. Неделина по специальности «Электроавтоматические системы и приборы подвижного и стационарного специального оборудования».

В 1974 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 140 научных публикаций и 20 авторских свидетельств на изобретения.

Область научных интересов — методы анализа и синтеза систем управления летательных аппаратов.

Эл. адрес: avgust.ar.@yandex.ru

**БАЛОНИН
Николай
Алексеевич**



Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 1982 году окончил Ленинградский электротехнический институт им. В. И. Ульянова (Ленина) по специальности «Автоматика и телемеханика».

В 2008 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором более 90 научных публикаций, в том числе трех монографий.

Область научных интересов — теория динамических систем, теория идентификации, теория операторов, теория матриц, вычислительные методы, интернет-робототехника, интернет-книжки с исполняемыми алгоритмами, научные социальные сети.

Эл. адрес: korbendfs@mail.ru

**БАРКОВСКИЙ
Евгений
Александрович**



Аспирант лаборатории информационных компьютерных технологий Института прикладных математических исследований Карельского научного центра РАН.

В 2012 году окончил Петрозаводский государственный университет по специализации «Информационные технологии в агропродовольственной сфере».

Является автором семи научных публикаций и одного учебного пособия.

Область научных интересов — задачи оптимального управления параллельными динамическими структурами данных, FIFO-очереди, деки с ограниченным входом.

Эл. адрес: barkevgen@gmail.com

**АРСЕНЬЕВ
Владимир
Николаевич**



Профессор кафедры бортовых информационных и измерительных комплексов Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург.

В 1978 году окончил Военно-инженерную Краснознаменную академию им. А. Ф. Можайского по специальности «Системы управления, электропроверочное и специальное оборудование летательных аппаратов».

В 1993 году защитил диссертацию на соискание ученой степени доктора технических наук.

Является автором 120 научных публикаций и двух патентов на изобретения.

Область научных интересов — методы анализа и синтеза систем управления летательных аппаратов.

Эл. адрес: vladar56@mail.ru

**БАРКОВ
Алексей
Владимирович**



Начальник сектора отдела системного программирования для космических аппаратов АО «Информационные спутниковые системы им. академика М. Ф. Решетнёва», Железнодорожск.

В 2002 году окончил Сибирский государственный аэрокосмический университет по специальности «Системы управления космическими аппаратами».

В 2006 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Область научных интересов — технологии программирования, системы автоматизации проектирования и сопровождения программного обеспечения.

Эл. адрес: barkov@reshenev.ru

**ВЕСЕЛОВ
Антон
Игорович**



Ассистент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2009 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Автоматизированные системы обработки информации и управления».

Является автором 26 научной публикации и шести патентов на изобретения.

Область научных интересов — теория информации, цифровая обработка изображений, распределенное кодирование источников, теория принятия решений.

Эл. адрес: anton.veselov@gmail.com

**ВОСТРИКОВ
Антон
Александрович**



Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2000 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети».

В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Является автором 45 научных публикаций и двух свидетельств о регистрации программного продукта.

Область научных интересов — распределенные и встраиваемые информационно-управляющие системы, обработка визуальной информации, опτικο-информационные системы.

Эл. адрес: vostricov@mail.ru

**ГИЛЬМУТДИНОВ
Марат
Равилевич**



Доцент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 1998 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Автоматизированные системы управления».

В 2002 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Является автором более 21 научной публикации и десяти патентов на изобретения.

Область научных интересов — теория информации (кодирование источников), методы цифровой обработки изображений и видеоданных, методы передачи видеoinформации.

Эл. адрес: mgilmutdinov@gmail.com

**ГРАНКИН
Максим
Андреевич**



Аспирант кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

В 2011 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации».

Является автором десяти научных публикаций и двух патентов на изобретения.

Область научных интересов — случайный множественный доступ, системы беспроводной связи, межмашинное взаимодействие.

Эл. адрес: m.a.grankin@gmail.com

**ГРИБКОВ
Алексей
Николаевич**



Доцент кафедры конструирования радиоэлектронных и микропроцессорных систем Тамбовского государственного технического университета.

В 2004 году окончил Тамбовский государственный технический университет по специальности «Проектирование и технология электронных средств».

В 2006 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Является автором 88 научных публикаций и двух патентов на изобретения.

Область научных интересов — теория оптимального управления многомерными объектами, методы разработки алгоритмического и программного обеспечения информационно-управляющих систем.

Эл. адрес: GribovAlexey@yandex.ru

**ГРИГОРЬЕВА
Ксения
Владимировна**



Доцент кафедры моделирования социально-экономических систем Санкт-Петербургского государственного университета.

В 2002 году окончила Санкт-Петербургский государственный университет по специальности «Прикладная математика».

В 2006 году защитила диссертацию на соискание ученой степени кандидата физико-математических наук.

Является автором 70 научных публикаций.

Область научных интересов — численные методы, негладкий анализ, математическая диагностика, теория оптимизации, теория игр и исследование операций, теория принятия решений.

Эл. адрес: k.grigorieva@spbu.ru, kсениа196247@mail.ru

**ГУРЕВИЧ
Виктор
Элизарович**



Профессор кафедры радиосистем и обработки сигналов Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича.

В 1961 году окончил Челябинский политехнический институт по специальности «Автоматические, телемеханические и электроизмерительные приборы и устройства».

В 1967 году защитил диссертацию на соискание ученой степени кандидата технических наук.

Является автором более 230 научных публикаций, 12 авторских свидетельств СССР на изобретения, двух патентов РФ на полезные модели.

Область научных интересов — общая теория связи, цифровые системы передачи информации, теория и техника автоматического управления.

Эл. адрес: gurvich23@mail.ru

**ДЖОКОВИЧ
Драгомир**



Гражданин Канады. Почетный профессор кафедры теоретической математики Университета Ватерлоо, Ватерлоо, Онтарио, Канада. В 1960 году окончил Белградский университет по специальности «Электротехника», Белград, Югославия. В 1963 году защитил диссертацию на соискание ученой степени доктора наук в Белградском университете. Является автором более 200 научных публикаций. Область научных интересов — линейная и полилинейная алгебра, теория групп, алгебра Ли и групп Ли, квантовая запутанность, комбинаторика. Эл. адрес: djokovic@uwaterloo.ca

**КОСМИНИНА
Наталья
Александровна**



Инженер-программист АО «Информационные спутниковые системы» им. академика М. Ф. Решетнёва, преподаватель филиала Сибирского федерального университета, Железнодорожный. В 2009 году окончила Сибирский федеральный университет по специальности «Автоматизированные системы обработки информации и управления», в 2011 году — по специальности «Информационные системы космических аппаратов и центров управления полетами». Является автором 12 научных публикаций и одного свидетельства о государственной регистрации программы для ЭВМ. Область научных интересов — теория языков программирования и методы трансляции. Эл. адрес: n.kosmylnina@list.ru

**КРАСИЛЬНИКОВА
Ольга
Ивановна**



Доцент кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 1978 году окончила Ленинградский институт авиационного приборостроения по специальности «Радиотехника». В 1991 году защитила диссертацию на соискание ученой степени кандидата технических наук. Является автором более 90 научных публикаций и семи учебных пособий. Область научных интересов — цифровая обработка изображений. Эл. адрес: OIKrasilnikova@yandex.ru

**ЕГОРОВ
Станислав
Геннадьевич**



Ассистент кафедры радиосистем и обработки сигналов Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича. В 2010 году окончил Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича по специальности «Радиотехника». Является автором восьми научных публикаций. Область научных интересов — цифровые системы передачи информации, многопозиционные радиолокационные системы, компьютерное моделирование в среде MatLab. Эл. адрес: sgegorov@gmail.com

**КРАСИЛЬНИКОВ
Николай
Николаевич**



Профессор кафедры информационно-сетевых технологий Санкт-Петербургского государственного университета аэрокосмического приборостроения, заслуженный деятель науки и техники РФ. В 1950 году окончил Ленинградский политехнический институт. В 1963 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 240 научных публикаций, в том числе пяти монографий. Область научных интересов — цифровая обработка изображений, статистическая теория передачи и восприятия изображений, математическое моделирование процессов обработки информации зрительной системой человека. Эл. адрес: NNKrasilnikov@yandex.ru

**КУРКИН
Илья
Александрович**



Ассистент кафедры конструирования радиоэлектронных и микропроцессорных систем Тамбовского государственного технического университета. В 2011 году окончил Тамбовский государственный технический университет по специальности «Проектирование и технология электронных средств». Является автором 28 научных публикаций. Область научных интересов — оптимальное управление, программные стратегии управления, многомерные объекты, методы искусственного интеллекта. Эл. адрес: allodsl@mail.ru

ЛАПИН
Александр
Анатолевич



Доцент межинститутской базовой кафедры прикладной физики и космических технологий Института космических и информационных технологий Сибирского федерального университета, начальник сектора АО «Информационные спутниковые системы» им. академика М. Ф. Решетнёва, Железногорск. В 1979 году окончил Новосибирский государственный университет по специальности «Прикладная математика». В 2005 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 32 научных публикаций. Область научных интересов — системы управления реального времени, методы проектирования и тестирования программного обеспечения. Эл. адрес: lapin_aa@mail.ru

ЛЯМИН
Андрей
Владимирович



Доцент кафедры компьютерных образовательных технологий, директор Центра дистанционного обучения Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики. В 1994 году окончил Санкт-Петербургский университет точной механики и оптики по специальности «Управление и информатика в технических системах». В 1997 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором более 145 научных публикаций. Область научных интересов — автоматизированные системы обработки информации и управления в образовании и науке. Эл. адрес: lyamin@mail.ifmo.ru

МУРОМЦЕВ
Дмитрий
Юрьевич



Профессор, заведующий кафедрой конструирования радиоэлектронных и микропроцессорных систем Тамбовского государственного технического университета. В 1996 году окончил Тамбовский государственный технический университет по специальности «Автоматизация технологических процессов и производств». В 2006 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 156 научных публикаций и десяти патентов на изобретения. Область научных интересов — анализ и синтез сложных систем на множестве состояний функционирования, энерго- и ресурсосбережение, энергосберегающие технологии, оптимальное управление и др. Эл. адрес: crems@crems.jesby.tstu.ru

ЛЕГАЛОВ
Александр
Иванович



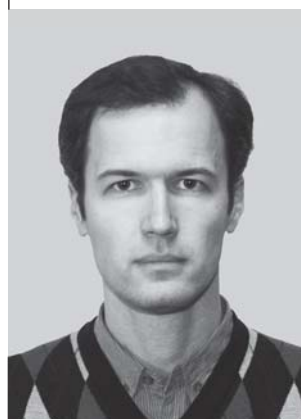
Профессор, заведующий кафедрой вычислительной техники Института космических и информационных технологий Сибирского федерального университета, Красноярск. В 1978 году окончил Красноярский политехнический институт по специальности «Конструирование радиоэлектронной аппаратуры». В 2005 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 170 научных публикаций и 12 свидетельств о регистрации программных продуктов. Область научных интересов — языки и парадигмы программирования, техника программирования, системное программирование и др. Эл. адрес: legalov@mail.ru

МАЛИЧЕНКО
Дмитрий
Александрович



Ведущий программист Института компьютерной безопасности вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2007 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации». Является автором девяти научных публикаций. Область научных интересов — помехоустойчивое кодирование, вычислительные сети, системы передачи и хранения данных. Эл. адрес: dml@vu.spb.ru

ПОНОМАРЕВ
Андрей
Васильевич



Старший научный сотрудник лаборатории интегрированных систем автоматизации Санкт-Петербургского института информатики и автоматизации РАН. В 2003 году окончил Тюменский государственный нефтегазовый университет по специальности «Автоматизированные системы обработки информации и управления». В 2012 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 30 научных публикаций. Область научных интересов — рекомендуемые системы, крауд-вычисления и краудсорсинг, машинное обучение, системы поддержки принятия решений. Эл. адрес: ponomarev@ias.spb.ru

СЕРГЕЕВ
Михаил
Борисович



Профессор, заведующий кафедрой вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения, директор НИИ информационно-управляющих систем Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики. В 1980 году окончил «ЛЭТИ». В 2001 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 100 научных публикаций и 14 патентов на изобретения. Область научных интересов — теория разрядных вычислений, методы проектирования специпроцессоров для систем контроля и управления и др.
Эл. адрес: mbse@mail.ru

СОКОЛОВ
Андрей
Владимирович



Ведущий научный сотрудник Института прикладных математических исследований Карельского научного центра РАН, профессор кафедры информатики и математического обеспечения Петрозаводского государственного университета. В 1974 году окончил Ленинградский государственный университет по специальности «Прикладная математика». В 2006 году защитил диссертацию на соискание ученой степени доктора физико-математических наук. Является автором более 100 научных публикаций. Область научных интересов — задачи оптимального управления динамическими структурами данных, динамическое распределение памяти, FIFO-очереди.
Эл. адрес: avs@krc.karelia.ru

ЧЕРЕПОВСКАЯ
Елена
Николаевна



Студентка кафедры компьютерных образовательных технологий факультета компьютерных технологий и управления, старший лаборант Центра дистанционного обучения Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики. Является автором пяти научных публикаций. Область научных интересов — разработка программных модулей с использованием сценарного языка программирования JavaScript, проведение и обработка экспериментов по анализу траектории взгляда.
Эл. адрес: cherepovskaya@cde.ifmo.ru

СИЛАНТЬЕВ
Сергей
Борисович



Профессор кафедры автономных систем управления Военно-космической академии им. А. Ф. Можайского, Санкт-Петербург. В 1984 году окончил Военный инженерный Краснознаменный институт им. А. Ф. Можайского по специальности «Системы управления летательных аппаратов». В 1997 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 60 научных публикаций и шести авторских свидетельств на изобретения. Область научных интересов — методы анализа и синтеза систем управления летательных аппаратов.
Эл. адрес: silantev2008@yandex.ru

ХИМЕНКО
Виталий
Иванович



Профессор кафедры компьютерной математики и программирования Санкт-Петербургского государственного университета аэрокосмического приборостроения, заслуженный деятель науки РФ, лауреат премии Правительства РФ в области науки и техники. В 1969 году окончил Куйбышевский политехнический институт по специальности «Приборные устройства». В 1992 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 150 научных публикаций. Область научных интересов — статистическая радиофизика, обработка информации в системах живой и неживой природы.
Эл. адрес: vih.13@yandex.ru

ЯСТРЕБОВ
Виктор
Анатольевич



Аспирант кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2013 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Комплексная защита объектов информатизации». Является автором шести научных публикаций и одного патента на изобретение. Область научных интересов — обработка сигналов, восстановление изображений, маскирование визуальных искажений, разреженное кодирование, компьютерное зрение, методы машинного обучения.
Эл. адрес: victor.yastrebov1@yandex.ru

