

ISSN 2713-3192
DOI 10.15622/ia.2022.21.1
<http://ia.spcras.ru>

ТОМ 21 № 1

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

INFORMATICS AND AUTOMATION



СПБ ФИЦ РАН

Санкт-Петербург
2022



INFORMATICS AND AUTOMATION

Volume 21 № 1, 2022

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

Founded in 2002

Founder and Publisher

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

Editor-in-Chief

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

Editorial Council

A. A. Ashimov	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
N. P. Veselkin	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
I. A. Kalyaev	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
Yu. A. Merkur'yev	Prof., Dr. Sci., Academician of the Latvian Academy of Sciences, Riga, Latvia
A. I. Rudskoi	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
V. Sgurev	Prof., Dr. Sci., Academician of the Bulgarian Academy of Sciences, Sofia, Bulgaria
B. Ya. Sovetov	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
V. A. Soyfer	Prof., Dr. Sci., Academician of RAS, Samara, Russia

Editorial Board

O. Yu. Gusikhin	Ph. D., Dearborn, USA
V. Delic	Prof., Dr. Sci., Novi Sad, Serbia
A. Dolgui	Prof., Dr. Sci., St. Etienne, France
M. N. Favorskaya	Prof., Dr. Sci., Krasnoyarsk, Russia
M. Zelezny	Assoc. Prof., Ph.D., Plzen, Czech Republic
H. Kaya	Assoc. Prof., Ph.D., Utrecht, Netherlands
A. A. Karpov	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
S. V. Kuleshov	Dr. Sci., St. Petersburg, Russia
A. D. Khomonenko	Prof., Dr. Sci., St. Petersburg, Russia
D. A. Ivanov	Prof., Dr. Habil., Berlin, Germany
K. P. Markov	Assoc. Prof., Ph.D., Aizu, Japan
R. V. Meshcheryakov	Prof., Dr. Sci., Moscow, Russia
N. A. Moldovian	Prof., Dr. Sci., St. Petersburg, Russia
V. Yu. Osipov	Prof., Dr. Sci., St. Petersburg, Russia
V. K. Pshikhopov	Prof., Dr. Sci., Taganrog, Russia
A. L. Ronzhin	Prof., Dr. Sci., Deputy Editor-in-Chief, St. Petersburg, Russia
H. Samani	Assoc. Prof., Ph.D., Plymouth, UK
V. Skormin	Prof., Ph.D., Binghamton, USA
A. V. Smirnov	Prof., Dr. Sci., St. Petersburg, Russia
B. V. Sokolov	Prof., Dr. Sci., St. Petersburg, Russia

Editor: A.S. Lopotova

Interpreter: Ya.N. Berezina

Art editor: N.A. Dormidontova

Editorial office address

14-th line V.O., 39, SPIIRAS, St. Petersburg, 199178, Russia,

e-mail: ia@spcras.ru, web: <http://ia.spcras.ru>

The journal is indexed in Scopus

The journal is published under the scientific-methodological supervision of Department for Nanotechnology and Information Technology of the Russian Academy of Sciences
© St. Petersburg Federal Research Center of the Russian Academy of Sciences, 2022

ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 21 № 1, 2022

Научный, научно-образовательный журнал с базовой специализацией
в области информатики, автоматизации, робототехники, прикладной математики
и междисциплинарных исследований.

Журнал основан в 2002 году

Учредитель и издатель

Федеральное государственное бюджетное учреждение науки
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
(СПб ФИЦ РАН)

Главный редактор

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

Редакционный совет

А. А. Ашимов	академик Национальной академии наук Республики Казахстан, д-р техн. наук, проф., Алматы, Казахстан
Н. П. Веселкин	академик РАН, д-р мед. наук, проф., Санкт-Петербург, РФ
И. А. Каляев	академик РАН, д-р техн. наук, проф., Таганрог, РФ
Ю. А. Меркурьев	академик Латвийской академии наук, д-р, проф., Рига, Латвия
А. И. Рудской	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. Сгурев	академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
Б. Я. Советов	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
В. А. Сойфер	академик РАН, д-р техн. наук, проф., Самара, РФ

Редакционная коллегия

О. Ю. Гусихин	д-р наук, Диаборн, США
В. Делич	д-р техн. наук, проф., Нови-Сад, Сербия
А. Б. Долгий	д-р наук, проф. Сент-Этьен, Франция
М. Железны	д-р наук, доцент, Пльзень, Чешская республика
Д. А. Иванов	д-р экон. наук, проф., Берлин, Германия
Х. Кайя	д-р наук, доцент, Утрехт, Нидерланды
А. А. Карпов	д-р техн. наук, доцент, Санкт-Петербург, РФ
С. В. Кулешов	д-р техн. наук, Санкт-Петербург, РФ
К. П. Марков	д-р наук, доцент, Аизу, Япония
Р. В. Мещеряков	д-р техн. наук, проф., Москва, РФ
Н. А. Молдовян	д-р техн. наук, проф., Санкт-Петербург, РФ
В.Ю. Осипов	д-р техн. наук, проф., Санкт-Петербург, РФ
В. Х. Пшихолов	д-р техн. наук, проф., Таганрог, РФ
А. Л. Ронжин	д-р техн. наук, проф., зам. главного редактора, Санкт-Петербург, РФ
Х. Самани	д-р наук, доцент, Плимут, Соединённое Королевство
В. А. Скормин	д-р наук, проф., Бингемптон, США
А. В. Смирнов	д-р техн. наук, проф., Санкт-Петербург, РФ
Б. В. Соколов	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. В. Уткин	д-р техн. наук, проф., Санкт-Петербург, РФ
М. Н. Фаворская	д-р техн. наук, проф., Красноярск, РФ
А. Д. Хомоненко	д-р техн. наук, проф., Санкт-Петербург, РФ
Л. Б. Шереметов	д-р техн. наук, Мехико, Мексика

Выпускающий редактор: А.С. Лопотова

Переводчик: Я.Н. Березина

Художественный редактор: Н.А. Дормидонтова

Адрес редакции

199178, г. Санкт-Петербург, 14-я линия В.О., д. 39

e-mail: ia@spcras.ru, сайт: <http://ia.spcras.ru>

Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук

© Федеральное государственное бюджетное учреждение науки

«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», 2022
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных в составе печатного периодического издания - журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ» статей по текущим экономическим, политическим, социальным и религиозным вопросам с обязательным указанием имени автора статьи и печатного периодического издания журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»

CONTENTS

Digital Information Telecommunication Technologies

K. Krotov

MATHEMATICAL MODEL AND ALGORITHM OF BRANCH AND
BOUNDARY METHOD FOR OPTIMIZING SOLUTIONS FOR PACKAGE
COMPOSITIONS IN MULTI-STAGE SYSTEMS 5

V. Zhmud, L. Dimitrov, G. Sablina, R. Hubert, J. Nosek, H. Wolfram
ON THE EXPEDIENCY AND POSSIBILITIES OF APPROXIMATING A
PURE DELAY LINK 41

S. Manaenko, S. Dvornikov, A. Pshenichnikov
THEORETICAL ASPECTS IN FORMING COMPLEX STRUCTURE SIGNAL 68

B. Mayorov
PROPERTIES OF HARMONIC AND COMPOSITE HALF-WAVES,
DETERMINATION OF THE UNIFORM TIME SAMPLING INTERVAL OF
DIGITAL SIGNAL PROCESSORS 95

Artificial Intelligence, Knowledge and Data Engineering

V. Yakimov, G. Maltsev

HYBRID NETWORK STRUCTURES AND THEIR USE IN DIAGNOSING
COMPLEX TECHNICAL SYSTEMS 126

G. Marshalko, J. Trufanova
POLYNOMIAL APPROXIMATIONS FOR SEVERAL NEURAL NETWORK
ACTIVATION FUNCTIONS 161

O. Fuentes, J. Savage, L. Contreras
A SLAM SYSTEM BASED ON HIDDEN MARKOV MODELS 181

СОДЕРЖАНИЕ

Цифровые информационно-телекоммуникационные технологии

К.В. Крогов

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И АЛГОРИТМ МЕТОДА ВЕТВЕЙ И ГРАНИЦ ДЛЯ ОПТИМИЗАЦИИ РЕШЕНИЙ ПО СОСТАВАМ ПАКЕТОВ В МНОГОСТАДИЙНЫХ СИСТЕМАХ 5

В.А. Жмудь, Л. Димитров, Г.В. Саблина, Р. Хьюберт, Я. Носек, Х. Вольфрам
О ЦЕЛЕСООБРАЗНОСТИ И ВОЗМОЖНОСТЯХ АППРОКСИМАЦИИ ЗВЕНА С ЧИСТЫМ ЗАПАЗДЫВАНИЕМ 41

С.С. Манаенко, С.В. Дворников, А.В. Пшеничников
ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ФОРМИРОВАНИЯ СИГНАЛЬНЫХ КОНСТРУКЦИЙ СЛОЖНОЙ СТРУКТУРЫ 68

Б.Г. Майоров
СВОЙСТВА ГАРМОНИЧЕСКИХ И СОСТАВНЫХ ПОЛУВОЛН, ОПРЕДЕЛЕНИЕ ИНТЕРВАЛА РАВНОМЕРНОЙ ДИСКРЕТИЗАЦИИ ВРЕМЕНИ ЦИФРОВЫХ СИГНАЛЬНЫХ ПРОЦЕССОРОВ 95

Искусственный интеллект, инженерия данных и знаний

В.Л. Якимов, Г.Н. Мальцев

ГИБРИДНЫЕ СЕТЕВЫЕ СТРУКТУРЫ И ИХ ИСПОЛЬЗОВАНИЕ ПРИ ДИАГНОСТИРОВАНИИ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ 126

Г.Б. Маршалко, Ю.А. Труфанова
ПОЛИНОМИАЛЬНЫЕ АППРОКСИМАЦИИ НЕКОТОРЫХ ФУНКЦИЙ АКТИВАЦИИ НЕЙРОННЫХ СЕТЕЙ 161

О. Фуэнтес, Х. Саваж, Л. Контрерас
СИСТЕМА SLAM, ОСНОВАННАЯ НА СКРЫТЫХ МАРКОВСКИХ МОДЕЛЯХ 181

К.В. КРОТОВ

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ И АЛГОРИТМ МЕТОДА ВЕТВЕЙ И ГРАНИЦ ДЛЯ ОПТИМИЗАЦИИ РЕШЕНИЙ ПО СОСТАВАМ ПАКЕТОВ В МНОГОСТАДИЙНЫХ СИСТЕМАХ

Кротов К.В. Математическая модель и алгоритм метода ветвей и границ для оптимизации решений по составам пакетов в многостадийных системах.

Аннотация. Современные методы решения задач планирования выполнения пакетов заданий в многостадийных системах характеризуются наличием ограничений на их размерность, невозможностью гарантированного получения лучших результатов в сравнении с фиксированными пакетами при различных значениях входных параметров задачи. В статье автором решена задача оптимизации составов пакетов заданий, выполняющихся в многостадийных системах, с использованием метода ветвей и границ. Проведены исследования различных способов формирования порядков выполнения пакетов заданий в многостадийных системах (эвристических правил упорядочивания пакетов заданий в последовательностях их выполнения на приборах МС). Определен способ упорядочивания пакетов в последовательностях их выполнения (эвристическое правило), обеспечивающий минимизацию общего времени реализации действий с ними на приборах. На основе полученного правила сформулирован способ упорядочивания типов заданий, в соответствии с которым их пакеты рассматриваются в процедуре метода ветвей и границ. Построена математическая модель процесса реализации действий с пакетами на приборах системы, которая обеспечивает вычисление его параметров. Выполнено построение метода формирования всех возможных решений по составам пакетов заданий для заданного их количества. Решения по составам пакетов заданий разных типов интерпретируются в процедуре метода ветвей и границ с целью построения оптимальной их комбинации. Для реализации метода ветвей и границ сформулирована процедура ветвления (разбиения), предполагающая формирование подмножеств решений, включающих пакеты разных составов заданий одного типа. Построены выражения для вычисления нижних и верхних оценок значений критерия оптимизации составов пакетов для сформированных в процедуре ветвления подмножеств. Процедура отсева предполагает исключение подмножеств, нижняя оценка которых не меньше рекорда. Для поиска оптимальных решений применена стратегия поиска в ширину, предусматривающая исследование всех подмножеств решений, включающих различные пакеты заданий одного типа, полученных в результате процедуры разбиения подмножеств заданий, не исключенных из рассмотрения после реализации процедуры отсева. Разработанные алгоритмы реализованы программно, что позволило получить результаты планирования выполнения пакетов заданий в многостадийной системе, являющиеся в среднем на 30 % лучшими, чем для фиксированных пакетов.

Ключевые слова: многостадийная система, пакеты заданий, метод ветвей и границ, эвристическое правило, расписания

1. Введение. Выполнение заданий разных типов, входящих в их наборы, в многостадийных системах (МС) предполагает формирование пакетов, количество заданий в которых определяется при оптимизации решений по их составам. Выполнение в МС заданий разных типов, включенных в составы пакетов, связано с переналадкой приборов МС с выполнения заданий одного типа на выполнение заданий другого

типа. Необходимость оптимизации составов пакетов вызвана неоднородностями длительностей выполнения заданий на приборах МС, а также неоднородностями длительностей переналадок приборов на выполнение заданий. Пакет – это совокупность заданий одного типа, выполняемых без переналадки приборов МС на выполнение заданий другого типа. Если в пакет включены все задания одного типа из их набора, такой пакет является фиксированным. Так как в МС выполняются ПЗ разных типов, тогда необходимо сформировать комплексные решения по составам ПЗ этих типов с учетом временных характеристик процесса реализации действий с ними в МС. Под оптимальным комплексным решением подразумевается комбинация составов ПЗ разных типов, обеспечивающая экстремум принятого показателя эффективности.

Современные методы планирования выполнения ПЗ в МС предполагают применение: целочисленного программирования (ЦЧП) для определения расписаний выполнения ПЗ [1]; метаэвристических алгоритмов (генетических алгоритмов (ГА), метода муравьиной колонии (МК)) при определении составов ПЗ и расписаний их выполнения [2-5]; эвристических правил при формировании составов ПЗ [6-8]; имитационного моделирования процесса обработки ПЗ в непрерывном производстве для оценки эффективности применения эвристик [9]. Математическая модель оптимизации решений при планировании выполнения ПЗ на параллельно функционирующих устройствах рассмотрена в [1]. Оптимизация решений с ее использованием предполагает определение значений переменных, соответствующих как составам пакетов, так и порядкам их выполнения на приборах МС. Применение данного подхода ограничено размерностью рассматриваемых задач. Применение ГА для решения задачи планирования ([2]) предполагает, что составы ПЗ определяются спросом на продукцию, построение расписания их выполнения реализуется с использованием этих алгоритмов. В [3,4] ГА применены для определения составов ПЗ (формируемые хромосомы соответствуют решениям по составам ПЗ). В [5] рассматривается метод планирования выполнения ПЗ на параллельных приборах на основе локального поиска совместно с методом МК при использовании эвристических правил. В силу стохастического характера рассмотренных алгоритмов при разных значениях входных параметров задачи они не гарантируют получение решений, приближающихся к оптимальным. В [6] эвристическое правило предусматривает включение в пакет заданий, длительности выполнения которых связаны с длительностью выполнения первого задания в этом пакете определенным образом. В [7] эвристические правила предусматривают форми-

рование множеств отклоняемых заданий и для каждого множества не отклоненных заданий формирование пакетов. В [8] метод формирования пакетов основывается на директивных сроках окончания выполнения заданий. В [9] определение составов ПЗ реализуется на основе имитационного моделирования. Определение маршрутов движения ПЗ, расписаний их выполнения реализуется с использованием правил, а оценивается с использованием имитационной модели. Отечественные авторы (Танаев В.С., Сотсков Ю.Н., Ковалев М.Я., Лазарев А.А., Кобак В.Г., Нейдорф Р.А.) рассматривают методы построения расписаний выполнения единичных заданий (ЕЗ), не объединяемых в пакеты. В [10] рассматриваются эвристические правила определения составов пакетов и методы построения расписаний выполнения фиксированных пакетов.

Метод ветвей и границ (МВГ) [11] является одним из базовых методов дискретной оптимизации решений. Однако его применение при планировании связано лишь с решением задач построения расписаний выполнения ЕЗ в обрабатывающих системах различного вида [12-20]. Работы [12,13] посвящены применению МВГ при построении расписаний выполнения ЕЗ на одном приборе (в [12] – с учетом директивных сроков окончания выполнения ЕЗ, в [13] – с учетом формирования партий, отгружаемых клиентам в соответствии со спросом на продукцию). В [14-15] рассматривается применение МВГ при построении расписаний выполнения ЕЗ в МС (без объединения их в пакеты). В [14] вводятся выражения для нижних оценок значений критерия оптимизации при формировании частичных последовательностей выполнения ЕЗ в МС, верхние оценки определяются на основе допустимых решений, получаемых с использованием метаэвристических алгоритмов. В [15] выражения для определения нижних и верхних оценок значений критерия формируются с учетом особенностей задачи: верхние оценки – с учетом простоев приборов, нижние – с учетом времени блокировки. В [16,17] МВГ применен при формировании расписаний выполнения ЕЗ на параллельных приборах. На каждом шаге процедуры ветвления МВГ в [16,17] формируются новые частичные последовательности выполнения ЕЗ на разных параллельно функционирующих приборах, соответствующие вершинам дерева. В [17] при формировании частичных расписаний учитывается отношение порядка для выполняемых заданий. Работы [18-20] посвящены использованию МВГ для решения задач построения расписаний выполнения проектов при ограничениях на ресурсы (RCPSP-задачи).

Определение составов ПЗ связано с решением задач ЦЧП, являющихся ограниченными по размерности, с применением метаэври-

стических алгоритмов и эвристических правил, приводящих к решениям, не гарантирующим приближения к оптимальным. Использование МВГ связано с построением расписаний выполнения ЕЗ на одном приборе, в МС, на параллельных приборах, построением расписаний выполнения проектов. Требуется развить МВГ для решения задач оптимизации составов ПЗ и построения расписаний их выполнения в МС.

2. Математическая модель процесса выполнения ПЗ в МС.

Для построения математической модели процесса выполнения ПЗ в МС, вида решений по составам ПЗ, расписаний выполнения ПЗ в МС в рассмотрении введены обозначения, представленные в Таблице 1.

Таблица 1. Обозначения параметров процесса выполнения ПЗ в МС

Обозначение	Назначение параметра
1	2
i	Идентификатор типа заданий ($i = \overline{1, n}$).
n^i	Количество заданий i -го типа ($i = \overline{1, n}$).
l	Номер прибора ($l = \overline{1, L}$).
m_i	Количество ПЗ i -го типа ($i = \overline{1, n}$) в решении по их составам.
M	Вектор количества пакетов i -ых типов ($i = \overline{1, n}$).
a_{ih}	Количество заданий i -го типа в h -ом пакете ($h = \overline{1, m_i}$).
A	Матрица составов ПЗ (размерность матрицы $n \times m$, где $m = \max_{i=1, n} (m_i)$).
π^l	Последовательность выполнения ПЗ на приборах МС.
n_p	Количество ПЗ в последовательностях π^l ($n_p = \sum_{i=1}^n m_i$).
j	Номер позиции ПЗ в последовательностях π^l их выполнения на приборах МС.
P	Матрица порядка выполнения ПЗ в π^l (размерность матрицы $n \times n_p$).
R	Матрица количества заданий i -ых типов в пакетах, занимающих в $\pi^l j$ -е позиции (размерность $n \times n_p$).
t_{li}	Длительность выполнения задания i -го типа на l -ом приборе МС ($i = \overline{1, n}$, $l = \overline{1, L}$).
T	Матрица длительностей выполнения заданий i -ых типов на приборах МС (размерность матрицы $L \times n$).

1	2
t_{ij}^l	Длительность переналадки l -го прибора с выполнения заданий i -го типа на выполнение заданий j -го типа.
T^l	Матрицы переналадки l -ых приборов ($l = \overline{1, L}$) (размерность каждой из матриц $n \times n$).
t_{ij}^{nl}	Моменты времени начала реализации действий с ПЗ i -го типа в j -й позиции в последовательности π^l ($l = \overline{1, L}$) их выполнения на приборах МС.
T^{nl}	Матрицы моментов времени начала выполнения ПЗ i -ых типов в j -ых позициях в последовательностях π^l ($l = \overline{1, L}$) (размерность матриц $n \times n_p$).
P_{ij}^l	Простои l -го прибора в ожидании готовности к выполнению ПЗ i -го типа, в j -ой позиции в последовательности π^l ($l = \overline{1, L}$) их выполнения на приборах МС.
N_1, N_2	Множества решений $[M, A]$ по составам ПЗ и расписаниям $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$ выполнения ПЗ в МС соответственно.

Порядок выполнения ПЗ на приборах МС одинаков, поэтому формируется одна матрица P и одна матрица R . Тогда расписание выполнения ПЗ на приборах МС имеет вид: $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$. В силу того, что расписание выполнения ПЗ на приборах МС формируется на основе решения $[M, A]$ путем интерпретации эвристического правила, тогда оно является результатом отображения множества решений по составам ПЗ вида $[M, A]$ на множество расписаний выполнения ПЗ, то есть $N_1 \rightarrow N_2([M, A])$.

Для формирования информационной модели задачи z планирования выполнения ПЗ каждому i -му типу заданий поставлен в соответствие кортеж вида [21]:

$$x_i = \langle i, T_i, \{T_i^l \mid l = \overline{1, L}\} \rangle, \quad (1)$$

где $T_i = (t_{1i}, t_{2i}, \dots, t_{Li})^T$ – i -й вектор-столбец длительностей выполнения заданий i -го типа на l -ых приборах МС (i -й столбец ($i = \overline{1, n}$) матрицы T длительностей выполнения заданий i -х типов на l -ых приборах МС), $T_i^l = (t_{i,1}^l, t_{i,2}^l, \dots, t_{i,n}^l)$ – i -я вектор-строка матрицы T^l ($l = \overline{1, L}$) длитель-

ностей переналадки l -го прибора с выполнения заданий i -го типа на выполнение заданий i' -го типа ($t_{i,i}^l$ соответствует времени первоначальной наладки l -го прибора на выполнение заданий i -го типа). Технологические маршруты (одинаковый фиксированный порядок приборов), дисциплины обслуживания (последовательное прохождение заданиями всех приборов в соответствии с маршрутами), правила выполнения для всех типов заданий одинаковы, поэтому эти параметры не вынесены в описание (1).

Так как в качестве входных параметров для задачи z заданы значения $n^i > 1$ ($i = \overline{1, n}$), то каждый тип заданий характеризуется кортежем вида $\langle x_i, n^i \rangle$ ($i = \overline{1, n}$). Тогда входными параметрами, на основе которых выполняется решение задачи z , является множество кортежей вида $\langle x_i, n^i \rangle$. Моменты времени поступления всех заданий в систему являются одинаковыми, равными 0, директивные сроки выполнения заданий не заданы, приоритеты выполнения заданий одинаковы, то эти параметры не вынесены в описание входных данных задачи. Множество входных данных обозначим через X . Вид множества X следующий: $X = \{ \langle x_i, n^i \rangle | i = \overline{1, n} \}$. Через Y обозначим множество результатов, формируемых системой планирования выполнения заданий при решении задачи z . В силу того, что ограничения на используемые ресурсы не заданы, тогда задача z , решение которой выполняется в системе планирования, характеризуется видом входной информации, видом критерия Kr , используемого при выборе лучшего из множества допустимых решений. Тогда решаемой в системе задаче z в соответствии поставлен кортеж параметров вида $z = \langle X, Y, Kr \rangle$.

В соответствие с [22] задача планирования выполнения ПЗ в МС представлена в виде: $\alpha | \beta | \gamma$, где α – тип обрабатывающей системы, β – тип решаемой задачи, γ – вид критерия оптимизации решений. Так как рассматривается многостадийная система – система *FlowShop*, то в качестве параметра α указывается F . Решение задачи связано с комплексным планированием выполнения ПЗ в МС (предусматривающим определение составов ПЗ всех типов и расписаний их выполнения в МС). Тогда в качестве параметра β указывается *batch* [10, 22]. Через n_p обозначено количество ПЗ в последовательностях π^l ($l = \overline{1, L}$) их выполнения на приборах МС ($n_p = \sum_{i=1}^n m_i$). В этом случае через t_{i,n_p}^{nL} обозначен момент времени начала выполнения последнего

n_p -го ПЗ некоторого i -го типа (занимающего последнюю n_p -ю позицию в последовательности π^L выполнения ПЗ на L -ом (последнем) приборе МС). Выражение $t_{i,n_p}^{nL} + r_{i,n_p} t_{L,i}$ определяет момент времени окончания выполнения этого ПЗ, занимающего n_p -ю позицию в последовательности π^L на L -ом приборе (то есть момент времени окончания выполнения всех ПЗ в МС).

Сформулированное эвристическое правило упорядочивания ПЗ в последовательностях π^l ($l = \overline{1, L}$) с учетом длительностей переналадок приборов является способом отображения множества решений N_1 на множество решений N_2 : $N_1 \rightarrow N_2([M, A])$. Тогда сформированное с использованием правила расписание $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$ не оптимизируется и соответствует решению по составам ПЗ $[M, A]$. Значение момента времени окончания выполнения ПЗ на приборах МС, определяемое выражением $t_{i,n_p}^{nL} + r_{i,n_p} t_{L,i}$, характеризует решение по составам ПЗ вида $[M, A]$. По этой причине оно используется в качестве критерия оптимизации решений по составам ПЗ (предусматривающих, в том числе, и построение расписаний выполнения ПЗ в МС). Тогда решаемая задача планирования имеет вид: $F \mid batch \mid (t_{i,n_p}^{nL} + r_{i,n_p} t_{L,i})$.

Входными данными задачи планирования являются: количество типов заданий n , количество заданий n^i ($i = \overline{1, n}$) каждого типа, матрицы длительностей выполнения заданий на приборах МС T и длительностей переналадок приборов МС T^l ($l = \overline{1, L}$). Результатами планирования является оптимальное решение по количествам и составам ПЗ разных типов вида $[M, A]$ и соответствующее ему расписание выполнения ПЗ в МС вида $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$, сформированное с использованием эвристического правила.

Задача оптимизации составов ПЗ решается в предположении, что длительности переналадок l -ых приборов ($l = \overline{1, L}$) с выполнения заданий i -го типа на выполнение заданий i' -го типа являются одинаковыми ($t_{ii'}^1 = t_{ii'}^2 = \dots = t_{ii'}^L$), длительности выполнения заданий i -х типов ($i = \overline{1, n}$) на l -ых приборах также одинаковы ($t_{li} = t_{2i} = \dots = t_{Li}$). Значения длительностей переналадок приборов с выполнения заданий i -го типа на выполнение заданий i' -го типа и длительностей выполнения заданий на приборах МС являются не убывающими. В соответствии с вве-

денными предположениями, связанными со значениями параметров t_{ii}^l и t_{ii} ($l = \overline{1, L}$), рассмотрены следующие эвристические правила, с использованием которых реализуется формирование порядка выполнения ПЗ в π^l ($l = \overline{1, L}$) на приборах МС:

- пакеты заданий с минимальной длительностью выполнения t_{ii} – в первую очередь (правило R1);
- пакеты заданий с максимальной длительностью выполнения t_{ii} – в первую очередь (правило R2);
- пакеты заданий с минимальной длительностью переналадки приборов t_{ii}^l – в первую очередь (правило R3);
- пакеты заданий с максимальной длительностью переналадки приборов t_{ii}^l – в первую очередь (правило R4).

Исследования указанных способов упорядочивания ПЗ в последовательностях π^l ($l = \overline{1, L}$) осуществлены путем построения диаграмм Ганта для каждого правила (при разных составах ПЗ). Для каждого эвристического правила и соответствующих ему последовательностей π^l ($l = \overline{1, L}$) выполнения ПЗ зафиксированы моменты времени окончания реализации действий с ними в МС. Анализ рассмотренных способов упорядочивания пакетов (с точки зрения минимизации окончания выполнения ПЗ в МС) показал, что использование параметра t_{ii}^l и правила R3 для построения расписаний позволяет получить минимальную длительность выполнения ПЗ. Реализация правила R3 упорядочивания ПЗ в последовательностях π^l ($l = \overline{1, L}$) предусматривает, что: 1) в первой позиции в последовательностях π^l ($l = \overline{1, L}$) размещается ПЗ с минимальным значением t_{ii}^l ; 2) во второй позиции в последовательностях π^l размещается ПЗ со значением $t_{i'i''}^l$, следующим за значением t_{ii}^l для заданий в пакете в первой позиции в π^l ($l = \overline{1, L}$); 3) в следующих позициях в π^l последовательно размещаются по одному пакету заданий других типов, длительности t_{ii}^l переналадки приборов с выполнения которых упорядочены по не убыванию (до n -го типа включительно); 4) в π^l размещаются по одному пакету заданий разных типов (от $(i=1)$ -го до $(i=n)$ -го типов), длительности t_{ii}^l переналадок приборов МС с выполнения которых упорядочены по не убыва-

нию; 5) действия, аналогичные пункту 4, продолжают до тех пор, пока в π^l ($l = \overline{1, L}$) не будут размещены все пакеты.

На основе сформулированного эвристического правила формирования порядка выполнения ПЗ в МС реализовано упорядочивание типов заданий с учетом значений t_{ii}^l . В соответствии с этим порядком типов заданий решения по составам их пакетов будут включаться с множества (подмножества) решений, формируемых при реализации процедуры разбиения (ветвления) в МВГ. В дополнение к обозначениям Таблицы 1 введены следующие обозначения: I – множество типов заданий, выполняемых в МС, упорядоченных с учетом значений t_{ii}^l ($|I| = n$); $f(t_{ii}^l)$ – способ упорядочивания элементов из I (связывания элементов из I отношением порядка \succ) в соответствии со значениями параметра t_{ii}^l . Способ $f(t_{ii}^l)$ предусматривает при $t_{ii}^l < t_{i'i''}^l$ связывание типов заданий i, i', i'' , входящих в I , отношением \succ следующим образом: $i' \succ i, i'' \succ i'$. Тогда ($i=I$)-ому типу заданий соответствует минимальное значение t_{ii}^l , n -му типу – максимальное t_{ii}^l .

Для построения математической модели процесса выполнения ПЗ в МС в рассмотрение введен параметр Pr_{ij}^l , характеризующий простой l -го прибора в ожидании начала выполнения ПЗ i -го типа, занимающего j -ю позицию в π^l ($i = \overline{1, n}, j = \overline{1, n_p}$). Простой ($l=I$)-го прибора в ожидании готовности ПЗ к выполнению отсутствуют. Данное утверждение представляется в следующем виде:

$$\text{Pr}_{11}^1 = \text{Pr}_{1,(n+1)}^1 = \dots = 0, \text{Pr}_{22}^1 = \text{Pr}_{2,(n+2)}^1 = \dots = 0, \dots, \text{Pr}_{n,n}^1 = \text{Pr}_{n,2n}^1 = \dots = 0. \quad (2)$$

Значения Pr_{ij}^l определяются на основе сформированной матрицы R следующим образом:

$$\text{Pr}_{i,j}^l = \max(0, (t_{i,j}^{nl-1} + r_{ij} t_{l-1,i}) - (t_{i',j-1}^{nl} + r_{i',j-1} t_{l,i'} + t_{i',i}^l)). \quad (3)$$

В (3) $t_{i,j}^{nl-1}$ – момент времени начала выполнения на ($l-1$)-м приборе ПЗ i -го типа, занимающего j -ю позицию в последовательности π^{l-1} , $t_{i',j-1}^{nl}$ – момент времени начала выполнения ПЗ i' -го типа, занимающего ($j-1$)-ю позицию в π^l . В соответствии с (2), (3) выражения для определения значений $t_{i,j}^{nl}$ имеют следующий вид:

– при $l=1$:

$$\begin{aligned}
 t_{11}^{n1} = 0; t_{22}^{n1} = t_{11}^{n1} + r_{11}t_{11} + t_{12}^1; \dots; t_{ij}^{n1} = t_{i-1,j-1}^{n1} + r_{i-1,j-1}t_{i-1,j-1} + t_{i-1,i}^1 \\
 t_{nn}^{n1} = t_{n-1,n-1}^{n1} + r_{n-1,n-1}t_{n-1,n-1} + t_{n-1,n}^1; \\
 t_{1,n+1}^{n1} = t_{n,n}^{n1} + r_{n,n}t_{n,n} + t_{n,1}^1; \dots; t_{n,2n}^{n1} = t_{n-1,2n-1}^{n1} + r_{n-1,2n-1}t_{n-1,n-1} + t_{n-1,n}^1;
 \end{aligned} \tag{4}$$

– при $l > 1$:

$$\begin{aligned}
 t_{11}^{nl} = \sum_{g=1}^{l-1} r_{11}t_{g1}; \\
 t_{i,j}^{nl} = t_{i,j-1}^{nl} + r_{i,j-1}t_{i,i}^l + t_{i,i}^l \text{ (при } j > 2).
 \end{aligned} \tag{5}$$

Выбор решения по составам ПЗ i -х типов ($i = \overline{1, n}$) осуществляется в соответствии с требованием минимизации окончания их выполнения в МС. Расписание $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$ является результатом отображения $N_1 \rightarrow N_2([M, A])$, соответствующего эвристическому правилу (матрицы P и R определяются в соответствии с эвристическим правилом на основе решения $[M, A]$, расписание не оптимизируется). Тогда решение $[M, A]$ может быть охарактеризовано моментом времени окончания выполнения ПЗ в МС. Выражение $t_{i,n_p}^{nL} + r_{i,n_p}t_{L,i}$ (где значение n_p ($n_p = \sum_{i=1}^n m_i$) и значение r_{i,n_p} (являющееся результатом отображения $f(t_{ii}^l)$) определяются в соответствии в решением $[M, A]$) введенное выше, определяет момент времени окончания выполнения в МС всех ПЗ. Тогда задача оптимизации составов ПЗ, выполняемых в МС, имеет следующий вид:

$$\min f([M, A], [P, R, \{T^{nl} \mid l = \overline{1, L}\}]), \tag{6}$$

где $f([M, A], [P, R, \{T^{nl} \mid l = \overline{1, L}\}]) = t_{i,n_p}^{nL} + r_{i,n_p}t_{L,i}$.

Ограничения на множества N_1 и $N_2([M, A])$ имеют вид:

– ограничение на количество заданий в пакетах:

$$\sum_{h=1}^{m_i} a_{ih} = n^i; \quad \sum_{i=1}^n \sum_{h=1}^{m_i} a_{ih} = \sum_{i=1}^n n^i; \quad (7)$$

– ограничение на количество ПЗ i -го типа в π^l ($l = \overline{1, L}$):

$$\sum_{j=1}^{n_p} p_{ij} = m_i; \quad (8)$$

– ограничение на общее количество ПЗ в π^l ($l = \overline{1, L}$):

$$\sum_{i=1}^n \sum_{j=1}^{n_p} p_{ij} = \sum_{i=1}^n m_i; \quad (9)$$

– ограничение на количество заданий i -го типа в пакетах в π^l :

$$\sum_{j=1}^{n_p} r_{ij} = \sum_{h=1}^{m_i} a_{ih}; \quad (10)$$

– ограничение на общее количество заданий i -ых типов ($i = \overline{1, n}$) в пакетах в π^l ($l = \overline{1, L}$), выполняемых на приборах МС:

$$\sum_{i=1}^n \sum_{j=1}^{n_p} r_{ij} = \sum_{i=1}^n \sum_{h=1}^{m_i} a_{ih}. \quad (11)$$

Постановка задачи оптимизации составов ПЗ, выполняемых в МС, сформулирована следующим образом. Заданными являются: количество n типов заданий; количество n^i ($i = \overline{1, n}$) заданий каждого типа; матрица T длительностей выполнения заданий в МС; матрицы T^l ($l = \overline{1, L}$) длительностей переналадок приборов. Требуется оптимизировать комплексное решение по составам ПЗ вида $[M, A]$ и сформировать соответствующее ему расписание $[P, R, \{T^{nl} \mid l = \overline{1, L}\}]$.

3. Применение метода вервей и границ при оптимизации комплексного решения по составам ПЗ, выполняемых в МС. Решение задачи оптимизации составов ПЗ обеспечивается построением: способа формирования решений по составам ПЗ i -х типов; способа

оптимизации комплексных решений по составам ПЗ разных типов (с использованием МВГ).

С целью определения множества решений по составам ПЗ и способа формирования решений по их составам введены условия [23]:

- количество заданий i -го типа в пакетах не менее 2 ($a_{ih} \geq 2$, $h = \overline{1, m_i}$); если при формировании решения по составам m_i ПЗ i -го типа для первого пакета получено $a_{i1} < 2$, то дальнейшее построение решений по составам ПЗ i -го типа в количестве m_i прекращается;

- формирование начального решения для количества m_i ПЗ i -го типа предполагает, что: $a_{ih} = 2$ ($h = \overline{2, m_i}$), $a_{i1} = n^i - \sum_{h=2}^{m_i} a_{ih}$;

- значения параметров m_i , задаваемые первоначально для заданий всех i -х типов ($i = \overline{1, n}$), равны 2 ($m_i = 2$);

- модификация количества ПЗ i -го типа предполагает, что параметр m_i увеличивается до тех пор, пока в начальном решении по составам m_i пакетов выполняется условие $a_{ih} \geq 2$ ($h = \overline{1, m_i}$); при $a_{i1} < 2$ формирование составов ПЗ i -го типа прекращается;

- формирование решений предполагает увеличение количества заданий в пакете с индексом $h' > 1$ и уменьшение количества заданий в пакете с $h = 1$; при условии $a_{i1} \geq a_{ih}$ ($h = \overline{2, m_i}$) модификация составов h -ых ПЗ продолжается; при условии $a_{i1} < a_{ih}$, решение по составам ПЗ в количестве m_i не рассматривается [23].

Ограничения множества решений по составам ПЗ имеют вид:

- 1) ограничение на количество ПЗ i -х типов ($i = \overline{1, n}$): $m_i \geq 2$;
 $m_i \leq \left\lceil \frac{n^i}{2} \right\rceil$, где $\lceil \cdot \rceil$ – операция округления в меньшую сторону;
- 2) ограничение на количество заданий i -х типов в пакетах ($i = \overline{1, n}$):

$$a_{i,h} \geq 2 \ (h = \overline{2, m_i}), \ a_{i,1} \geq 2, \ a_{i,1} = n^i - \sum_{h=2}^{m_i} a_{i,h}, \ a_{i,1} \geq a_{i,h} \ (h = \overline{2, m_i});$$

$$a_{i,1} \leq n^i - 2(m_i - 1); \ a_{i,h} \leq \left\lceil \frac{n^i - 2(m_i - 2)}{2} \right\rceil;$$

3) ограничение на значения количества заданий i -х типов в пакетах:

$$a_{i,1} \in \{2, 3, \dots, n^i - 2(m_i - 1)\}, \quad a_{i,h} \in \left\{2, 3, \dots, \left\lceil \frac{n^i - 2(m_i - 2)}{2} \right\rceil\right\} \quad (h = \overline{2, m_i}).$$

Для оптимизации решения по составам ПЗ с использованием МВГ в рассмотрение введены обозначения, приведенные в Таблице 2.

Значение параметра m_i^{\max} , соответствующего максимальному количеству ПЗ i -го типа, определяется следующим образом:

$$m_i^{\max} = \left\lceil \frac{n^i}{2} \right\rceil. \quad \text{В матрице } A_i^n \text{ хранятся все возможные решения по составам ПЗ } i\text{-го типа, сформированные в соответствии с рассмотренным ниже алгоритмом.}$$

Значение параметра $r_{A_i^n}$ соответствует количеству строк в матрице A_i^n , определяется выражением вида:

$$r_{A_i^n} = \sum_{m_i=2}^{m_i^{\max}} \left\lceil \frac{n^i + 1}{m_i} \right\rceil, \quad \text{где } \lceil \cdot \rceil - \text{ округление в большую сторону.}$$

Матрицы $A_{1i}^{\text{быв}}$ и $A_{2i}^{\text{быв}}$ предназначены для временного хранения решений по составам ПЗ i -го типа ($i = \overline{1, n}$): в $A_{1i}^{\text{быв}}$ хранятся решения, на основе которых формируются новые решения, записываемые в матрицу $A_{2i}^{\text{быв}}$.

Способ определения составов ПЗ i -го типа формирует на основе решений из $q1_i$ -х строк матрицы $A_{1i}^{\text{быв}}$ новые решения, которые сохраняются в матрице $A_{2i}^{\text{быв}}$. После того, как новые решения сформированы, решения из матрицы $A_{1i}^{\text{быв}}$ переписываются в матрицу A_i^n , а решения из матрицы $A_{2i}^{\text{быв}}$ – в $A_{1i}^{\text{быв}}$ для формирования новых решений. Значения $r_{A_{1i}^{\text{быв}}}$ и $r_{A_{2i}^{\text{быв}}}$ определяются выражением

$$r_{A_{1i}^{\text{быв}}} = r_{A_{2i}^{\text{быв}}} = \max_{2 \leq m_i \leq m_i^{\max}} \left(\left\lceil \frac{n^i + 1}{m_i} \right\rceil \right).$$

В соответствии с разработанным методом формирования составов ПЗ в каждой матрице A_i^n ($i = \overline{1, n}$) накапливаются все решения по составам пакетов i -го типа. Матрицы $A_{1i}^{\text{быв}}$ и $A_{2i}^{\text{быв}}$ используются для формирования новых решений по составам ПЗ i -го типа.

Таблица 2. Обозначения, используемые при формировании комплексного решения по составам ПЗ i -ых типов, выполняемых в МС

Обозначение	Назначение
1	2
A_i^n	Матрицы, предназначенные для хранения решений по составам ПЗ i -ых типов, размерность $r_{A_i^n} \times m_i^{\max}$.
$r_{A_i^n}$	Количество строк в матрице A_i^n .
m_i^{\max}	Максимальное количество ПЗ i -го типа.
q_i	Индекс строки в матрице A_i^n ($i = \overline{1, n}$).
$(a_{q_i, h})_i^n$	Количество заданий i -го типа в ($h=1$)-ом пакете для q_i -го решения по составам пакетов этого типа.
n_p^i	Счетчик количества решений по составам ПЗ i -го типа в матрице A_i^n ($i = \overline{1, n}$).
$A_{1i}^{\text{бывф}}, A_{2i}^{\text{бывф}}$	Матрицы, предназначенные для хранения решений по составам ПЗ i -го типа при их формировании (размерность матриц $r_{A_{1i}^{\text{бывф}}} \times m_i^{\max}$ и $r_{A_{2i}^{\text{бывф}}} \times m_i^{\max}$)
$r_{A_{1i}^{\text{бывф}}}, r_{A_{2i}^{\text{бывф}}}$	Количество строк в матрицах $A_{1i}^{\text{бывф}}, A_{2i}^{\text{бывф}}$.
$q1_i, q2_i$	Индексы строк в матрицах $A_{1i}^{\text{бывф}}$ и $A_{2i}^{\text{бывф}}$.
n_{p1}^i, n_{p2}^i	Счетчики решений по составам ПЗ i -ых типов, размещаемых в матрицах $A_{1i}^{\text{бывф}}$ и $A_{2i}^{\text{бывф}}$.
G_{iq_i}	Множество решений по составам ПЗ, содержащее решение по составу ($h=1$)-го ПЗ i -го типа, соответствующее q_i -ой строке матрицы A_i^n .
$I_{МВГ}$	Множество типов заданий, для которых при реализации МВГ выполнено формирование множеств G_{iq_i} .

На основе сформулированного способа определения составов ПЗ, в рассмотрение введены условия, позволяющие сократить количество формируемых решений [23]. Если для текущего рассматриваемого решения ($q1_i$ -й строки матрицы $A_{1i}^{\text{бывф}}$) выполняется условие $(a_{q1_i, h})_{1i}^{\text{бывф}} = (a_{q1_i, h+1})_{1i}^{\text{бывф}}$, то увеличение на 1 значения $(a_{q1_i, h})_{1i}^{\text{бывф}}$ (при неизменном значении $(a_{q1_i, h+1})_{1i}^{\text{бывф}}$), а затем увеличение на 1 значения $(a_{q1_i, h+1})_{1i}^{\text{бывф}}$ (при неизменном значении $(a_{q1_i, h})_{1i}^{\text{бывф}}$) обуславливает получение одинаковых решений по составам ПЗ i -го типа.

На основе этого условия сформулировано следующее утверждение, позволяющее ограничить количество решений [23]: если при формировании решения по составам m_i ПЗ i -го типа на основе решения, соответствующего $q_1 i$ -й строке матрицы $A_{li}^{\text{бывф}}$, для элементов этой строки выполняется $(a_{q_1, h}^{\text{бывф}})_{li} = (a_{q_1, h+1}^{\text{бывф}})_{li}$, то не требуется увеличивать значение элемента $(a_{q_1, h+1}^{\text{бывф}})_{li}$; требуется выполнить переход к элементу $(a_{q_1, h+j}^{\text{бывф}})_{li}$ ($(h+j) \leq m_i$), такому, что $(a_{q_1, h}^{\text{бывф}})_{li} \neq (a_{q_1, h+j}^{\text{бывф}})_{li}$, значение которого будет изменено. Если $(a_{q_1, h}^{\text{бывф}})_{li} = (a_{q_1, h+1}^{\text{бывф}})_{li}$ (в общем виде $(a_{q_1, h}^{\text{бывф}})_{li} = (a_{q_1, h+j}^{\text{бывф}})_{li}$), то на основе решения, представленного $q_1 i$ -й строкой матрицы $A_{li}^{\text{бывф}}$, будет получено решение, дублирующее сформированное ранее. В том случае, если для сформированного решения по составам m_i ПЗ i -го типа выполняется условие $a_{q_2, 1} < a_{q_2, h}$ ($h = \overline{2, m_i}$), то это решение дублирует сформированное ранее, оно не рассматривается и не используется для последующего формирования новых решений [23].

Для реализации алгоритма построения решений по составам ПЗ i -го типа необходимо сформировать начальное решение для $m_i = 2$ (задать составы ПЗ (при $q_1 i = 1$ и $n_{p1}^i = 1$): $a_{q_1, h} = 2$ ($h = \overline{2, m_i}$),

$a_{q_1, 1} = n^i - \sum_{h=2}^{m_i} a_{q_1, h}$; если $a_{q_1, 1} \geq 2$, то сформировано корректное решение по составам $m_i = 2$ ПЗ i -го типа. Если $a_{q_1, 1} < 2$, то для i -го типа заданий формируется фиксированный пакет, состав которого сохраняется в матрице A_i^n : $q_i = 1, a_{q_i, 1} = n^i, n_p^i = 1$.

Входными данными для алгоритма формирования решений по составам ПЗ i -го типа на основе решений, находящихся в матрице $A_{li}^{\text{бывф}}$, являются: n_{p1}^i – количество решений по составам ПЗ i -го типа в матрице $A_{li}^{\text{бывф}}$ ($n_{p1}^i = 1$ в начальном решении при $m_i = 2$), матрица $A_{li}^{\text{бывф}}$. Алгоритм построения решений по составам ПЗ i -го типа на основе решений, находящихся в матрице $A_{li}^{\text{бывф}}$, имеет порядок шагов:

1. Задать значения параметров: $q_1 i = 1, n_{p2}^i = 0, q_2 i = 0, n_p^i = 0$.

2. Задать номер пакета h' в решении, соответствующем $q1_i$ -й строке матрицы $A_i^{\delta y \phi}$, состав которого изменяется, равным 2 ($h' = 2$).

3. Модифицировать значение $q2_i$: $q2_i = q2_i + 1$; в h' -м пакете увеличить количество заданий на 1 – сформировать решение по составам ПЗ i -го типа ($q2_i$ -ю строку матрицы $A_{2i}^{\delta y \phi}$): $(a_{q2_i, h})_{2i}^{\delta y \phi} = (a_{q1_i, h})_{1i}^{\delta y \phi}$ ($h = \overline{2, h'-1} \& h = \overline{h'+1, m_i}$), $(a_{q2_i, h'})_{2i}^{\delta y \phi} = (a_{q1_i, h'})_{1i}^{\delta y \phi} + 1$,

$$(a_{q2_i, 1})_{2i}^{\delta y \phi} = n^i - \sum_{h=1}^{m_i} (a_{q1_i, h})_{1i}^{\delta y \phi}; n_{p2}^i = n_{p2}^i + 1.$$

4. Проверить условие $(a_{q2_i, 1})_{2i}^{\delta y \phi} < (a_{q2_i, h'})_{2i}^{\delta y \phi}$. В случае его выполнения сформированное в виде $q2_i$ -й строки матрицы $A_{2i}^{\delta y \phi}$ решение не интерпретируется, тогда $n_{p2}^i = n_{p2}^i - 1$ и $q2_i = q2_i - 1$; если $n_{p2}^i = 0$, то перейти на пункт 9; если $n_{p2}^i > 0$, то перейти на пункт 8; при $(a_{q2_i, 1})_{2i}^{\delta y \phi} \geq (a_{q2_i, h'})_{2i}^{\delta y \phi}$ полученное решение сохранено в матрице $A_{2i}^{\delta y \phi}$, выполнить задание значения параметра $j=1$ (j -шаг изменения номера пакета h'), перейти на пункт 5.

5. Если $(h' + j) \leq m_i$, то проверить $(a_{q1_i, h'})_{1i}^{\delta y \phi} > (a_{q1_i, h'+j})_{1i}^{\delta y \phi}$; если условие не выполняется, то перейти на пункт 6. Если условие выполняется, то $h' = h' + j$, перейти на пункт 3. При $(h' + j) > m_i$ перейти на пункт 7.

6. Модифицировать параметр $j=j+1$. Перейти на пункт 5.

7. При $(h' + j) > m_i$ сформированы все решения (размещенные в матрице $A_{2i}^{\delta y \phi}$) с использованием решения из $q1_i$ -ой строки матрицы $A_{1i}^{\delta y \phi}$; модифицировать значение $q1_i$: $q1_i = q1_i + 1$; если $q1_i \leq n_{p1}^i$, то перейти на пункт 2; если $q1_i > n_{p1}^i$, то перейти на пункт 8.

8. Если $n_{p2}^i > 1$, то выполнить сравнение решений, хранимых в матрице $A_{2i}^{\delta y \phi}$, с точки зрения дублирования ими друг друга (сравнение предполагает формирование копии матрицы $A_{2i}^{\delta y \phi}$, в копии матрицы $A_{2i}^{\delta y \phi}$ – упорядочивание элементов каждой $q2_i$ -й строке ($q2_i = \overline{1, n_{p2}^i}$) по не возрастанию значений, поэлементное сравнение каждой $q2_i$ -й строки ($q2_i = \overline{1, n_{p2}^i} - 1$) с другими $q2'_i$ -ми строками

($q2'_i = q2_i + 1, n_{p2}^i$), удаление из матрицы $A_{2i}^{\delta y\phi}$ строк с индексами $q2'_i$, которые соответствуют строкам в копии матрицы $A_{2i}^{\delta y\phi}$, дублирующим рассматриваемую $q2_i$ -ю строку; при удалении $q2'_i$ -й строки из матрицы $A_{2i}^{\delta y\phi}$ реализуется изменение $n_{p2}^i = n_{p2}^i - 1$, переупорядочивание решений в $A_{2i}^{\delta y\phi}$; если $n_{p2}^i = 1$, то перейти на пункт 9.

9. Перезаписать решения из матрицы $A_{1i}^{\delta y\phi}$ в матрицу A_i^n : $(a_{q_i, h})_i^n = (a_{q_i, h})_{1i}^{\delta y\phi}$ при $h = \overline{1, m_i}$, $q1_i = \overline{1, n_{p1}^i}$, $q_i = \overline{n_p^i + 1, n_p^i + n_{p1}^i}$; модификация значений n_{p1}^i и n_p^i : $n_p^i = n_p^i + n_{p1}^i$, $n_{p1}^i = 0$; если $n_{p2}^i > 0$, то осуществить перезапись решений из матрицы $A_{2i}^{\delta y\phi}$ в матрицу $A_i^{\delta y\phi}$: $(a_{q_i, h})_{1i}^{\delta y\phi} = (a_{q_i, h})_{2i}^{\delta y\phi}$ при $h = \overline{1, m_i}$, $q1_i = \overline{1, n_{p2}^i}$, $q2_i = \overline{1, n_{p2}^i}$; задать значения параметров алгоритма: $q1_i = 1$, $n_{p1}^i = n_{p2}^i$, $q2_i = 0$, $n_{p2}^i = 0$; перейти на пункт 2.

10. Если $n_{p2}^i = 0$, то реализовать модификацию значения параметра m_i (увеличить количества ПЗ i -го типа): $m_i = m_i + 1$; сформировать начальное решение по составам ПЗ: $q1_i = 1$, $a_{q_i, h} = 2$ ($h = \overline{2, m_i}$), $a_{q1, 1} = n^i - \sum_{h=2}^{m_i} a_{q_i, h}$; $n_{p1}^i = 1$; если $a_{q1, 1} \geq 2$, то инициализировать параметры алгоритма: $q2_i = 0$, $n_{p2}^i = 0$; перейти на пункт 2; если $a_{q1, 1} < 2$, то начальное решение по составам ПЗ i -го типа не является корректным, перейти на пункт 11.

11. Останов алгоритма.

В результате реализации алгоритма каждая q_i -я строка матрицы A_i^n ($q_i = \overline{1, n_p^i}$) является решением по составам ПЗ i -го типа.

Для применения МВГ при решении задачи оптимизации составов ПЗ n типов построена математическая модель процесса их выполнения в МС. С использованием этой модели на основе матрицы R , соответствующей порядкам выполнения ПЗ из q_i -х строк матриц A_i^n в π^l , определяются значения t_{ij}^{nl} для ПЗ i -х типов, занимающих j -е позиции в π^l . Матрица R определяется для комбинации решений по со-

ставам ПЗ i -х типов из q_i -х строк ($q_i = \overline{1, n_p^i}$) матриц A_i^n (значения $r_{i,j}$ элементов R определяются на основе значений элементов $a_{q_i,h}$ из q_i -х строк матриц A_i^n в соответствии с отображением $f(t_{ii}^l)$).

Для использования МВГ [11] с целью оптимизации комплексно-решения по составам ПЗ сформулированы:

- способ разбиения множеств решений G_{iq_i} , содержащих ПЗ i -го типа, на подмножества решений $G_{(i+1)q_{(i+1)}}$, включающих ПЗ $(i+1)$ -го типа ($q_{(i+1)} = \overline{1, n_p^{i+1}}$) (процедура ветвления);

- способ вычисления верхней оценки (рекорда) H^6 и способ вычисления нижних оценок H_{i,q_i}^n для подмножеств G_{iq_i} , включающих решения по составам ПЗ i -го типа из q_i -х строк матриц A_i^n .

Применение МВГ предусматривает, что комплексное решение по составам ПЗ разных типов (дерево решений) формируется последовательно. Первоначально рассматривается множество решений G_{1q_1} , содержащее пакет заданий $(i=1)$ -го типа из q_1 -го решения в матрице A_1^n (корневая вершина дерева МВГ). В результате разбиения G_{1q_1}

формируются подмножества решений G_{2q_2} ($q_2 = \overline{1, n_p^2}$), включающие (наряду с ПЗ $(i=1)$ -го типа) ПЗ $(i=2)$ -го типа. Аналогичным образом будет реализовываться процедура разбиения для каждого множества, которое не исключено из рассмотрения после реализации процедуры отсева. В соответствии с изложенной процедурой будет сформировано n_p^l деревьев. В корне каждого находится вершина, соответствующая

$(h=1)$ -у ПЗ $(i=1)$ -го типа в q_1 -м решении ($q_1 = \overline{1, n_p^1}$) по составам ПЗ из матрицы A_1^n . Выбор лучшего решения по составам ПЗ осуществляется путем сравнения значений критерия $f(\cdot)$ для решений, полученных при реализации МВГ для каждого дерева.

В матрице R определяются элементы, которые соответствуют ПЗ, размещенным в π^l при реализации текущей и предыдущих итераций процедуры разбиения (ветвления) в МВГ. Алгоритм формирования

ния матрицы R на основе значений $a_{q_i, h}$ из q_i -х строк ($q_i = \overline{1, n_p^i}$) матриц A_i^n активизируется каждый раз при реализации процедуры разбиения подмножеств для каждого нового подмножества. Сформированная матрица R используется при вычислении нижних оценок значений критерия $f(\cdot)$ для каждого из сформированных подмножеств, вычисления значений критерия $f(\cdot)$ для допустимых решений по порядку выполнения ПЗ, соответствующих рассмотренным подмножествам, и фиксированных пакетов для i -х типов заданий, соответствующих не сформированным подмножествам (при модификации рекорда [11]).

Реализация процедуры ветвления предусматривает, что каждой вершине соответствует кортеж параметров вида $\langle i, k_s, q_i, h_i \rangle$, где k_s – номер вершины, которая внесена в формируемое дерево решений МВГ, q_i – номер решения (индекс строки в матрице A_i^n), которое сопоставлено с рассматриваемой вершиной, h_i – номер ПЗ i -го типа из решения, соответствующего q_i -й строке матрицы A_i^n . Корневой вершине соответствует набор вида $\langle 1, 1, q_i, 1 \rangle$.

Для реализации МВГ определены три базовые процедуры [11]: процедуру поиска, процедуру отсева и процедуру ветвления. В качестве процедуры поиска использован поиск в ширину (то есть исследуются все подмножества решений G_{iq_i} ($q_i = \overline{1, n_p^i}$), включающие $(h=1)$ -е ПЗ разных составов рассматриваемого i -го типа (из q_i -х строк матрицы A_i^n ($q_i = \overline{1, n_p^i}$)), сформированные на текущей итерации МВГ при реализации процедуры ветвления). Исследование сформированных в результате разбиения подмножеств G_{iq_i} предполагает вычисление оценок H_{i, q_i}^h для каждого из них, вычисление значений оценок H_s^6 критерия $f(\cdot)$ для формируемых допустимых решений (с целью обновления рекорда при условии $H_s^6 < H^6$).

Процедура разбиения предполагает, что каждому подмножеству $G_{(i-1)q_{(i-1)}}$, содержащему решение по составам ПЗ из q_{i-1} -й строки матрицы A_{i-1}^n ($q_{i-1} = \overline{1, n_p^{i-1}}$), не исключенному из рассмотрения по-

сле отсева, ставятся в соответствие подмножества G_{iq_i} , содержащие q_i -е решения по составам ПЗ i -го типа ($q_i = \overline{1, n_p^i}$). Каждому подмножеству решений G_{iq_i} ($q_i = \overline{1, n_p^i}$) соответствует вершина дерева МВГ с кортежем $\langle i, k_s, q_i, h_i \rangle$. При реализации обхода дерева МВГ в глубину определяются идентификаторы решений q_i , пакеты из которых размещаются в последовательностях π^l . В результате формируется матрица R , на основе которой с использованием модели процесса выполнения ПЗ в МС реализуется вычисление значений H_{i,q_i}^H . Процедура отсева предполагает, что из рассмотрения исключаются множества G_{iq_i} , для которых $H_{i,q_i}^H \geq H^e$. Вычисление верхней оценки значений $f(\cdot)$ для множества решений G_{lq_l} осуществляется на основе сформированного в соответствии с введенным правилом расписания выполнения фиксированных ПЗ i -ых типов. Введено условие: $Pr_1^l = Pr_2^l = \dots = Pr_n^l = 0, Pr_l^l = 0 (l = \overline{1, L})$. Простой ($l=2$)-го прибора в ожидании готовности фиксированных пакетов ($i=2$)-го и ($i=3$)-го типов определяются в соответствии с выражениями вида:

$$Pr_2^2 = \max(0, (n^1 t_{11} + t_{12}^1 + n^2 t_{12}) - (\sum_{l=1}^2 n^l t_{l,1} + t_{12}^2));$$

$$Pr_3^2 = \max(0, (n^1 t_{11} + t_{12}^1 + n^2 t_{12} + t_{23}^1 + n^3 t_{13}) - (\sum_{l=1}^2 n^l t_{l,1} + t_{12}^2 + Pr_2^2 + n^2 t_{22} + t_{23}^2)).$$

В общем виде формула для вычисления простоев ($l=2$)-го прибора в ожидании готовности фиксированного ПЗ i -го типа к выполнению имеет вид:

$$Pr_i^2 = \max(0, (\sum_{j=1}^i n^j t_{1j} + \sum_{j=1}^{i-1} t_{j,j+1}^1) - (n^1 t_{11} + \sum_{j=1}^{i-1} n^j t_{2j} + \sum_{j=1}^{i-1} Pr_j^2 + \sum_{j=1}^{i-1} t_{j,j+1}^1)).$$

Выражение для простоев ($l=3$)-го прибора в ожидании начала выполнения фиксированных пакетов ($i=2$)-го и ($i=3$)-го типов:

$$Pr_2^3 = \max(0, (\sum_{l=1}^2 n^l t_{l1} + t_{l2}^2 + Pr_2^2 + n^2 t_{22}) - (\sum_{l=1}^3 n^l t_{l1} + t_{l2}^3));$$

$$Pr_3^3 = \max(0, (\sum_{l=1}^2 n^l t_{l1} + t_{l2}^2 + Pr_2^2 + n^2 t_{22} + t_{23}^2 + Pr_3^2 + n^3 t_{23}) -$$

$$- (\sum_{l=1}^3 n^l t_{l1} + t_{l2}^3 + Pr_2^3 + n^2 t_{32} + t_{23}^3)).$$

Общей вид формулы для вычисления простоев l -го прибора в ожидании готовности фиксированного ПЗ i -го типа к выполнению:

$$Pr_i^l = \max(0, (\sum_{q=1}^{l-2} n^l t_{q,l} + \sum_{j=1}^i n^j t_{l-1,j} + \sum_{j=1}^i Pr_j^{l-1} + \sum_{j=1}^{i-1} t_{j,j+1}^{l-1}) -$$

$$- (\sum_{q=1}^{l-1} n^l t_{q,l} + \sum_{j=1}^{i-1} n^j t_{l,j} + \sum_{j=1}^{i-1} Pr_j^l + \sum_{j=1}^{i-1} t_{j,j+1}^l)).$$

Выражение для вычисления верхней оценки (рекорда) в корне:

$$H^6 = \sum_{l=1}^L n^l t_{l,l} + t_{l2}^L + Pr_2^L + n^2 t_{L,2} + t_{23}^L + Pr_3^L + n^3 t_{L,3} + ..$$

$$.. + Pr_{n-1}^L + n^{(n-1)} t_{L,(n-1)} + t_{(n-1),n}^L + Pr_n^L + n^n t_{L,n}. \quad (12)$$

Выражение (12) представлено в общем виде:

$$H^6 = \sum_{l=1}^{L-1} n^l t_{l,l} + \sum_{j=1}^n n^j t_{L,j} + \sum_{j=1}^n Pr_j^L + \sum_{j=1}^{n-1} t_{j,j+1}^L. \quad (13)$$

Определение нижней оценки $H_{iq_i}^H$ осуществляется при формировании множества решений G_{iq_i} , содержащего пакет i -го типа из q_i -й строки (решения) матрицы A_i^n . Формирование множества G_{iq_i} сопровождается добавлением в последовательности π^l ($l = \overline{1, L}$) пакета i -го типа, параметры которого соответствуют вершине дерева МВГ.

Нижняя оценка $H_{iq_i}^H$ для множества G_{iq_i} , полученного в результате ветвления, формируется с учетом добавленных в π^l ПЗ рассмотренных i' -ых типов ($i' = \overline{1, i}$), размещения в π^L фиксированных ПЗ других типов, для решений по составам которых вершины в дереве не сформированы, размещения в π^L «оставшихся» заданий i' -х типов ($i' = \overline{1, i}$).

Для вычисления значения $H_{2q_2}^H$ синтезируется матрица R (с учетом правила формирования последовательности π^L), на основе которой реализуется вычисление значений t_{ij}^{nl} ($i = \overline{1, 2}$, $j = \overline{1, 2}$, $l = \overline{1, L}$). Оценка $H_{2q_2}^H$ определяется следующим образом:

$$H_{2q_2}^H = t_{22}^{nL} + r_{22}t_{L,2} + t_{23}^L + n^3t_{L,3} + t_{34}^L + n^4t_{L,4} + .. \\ .. + t_{(n-1),n}^L + n^nt_{L,n} + t_{n,1}^L + (n^l - r_{11})t_{L1} + t_{1,2}^L + (n^l - r_{22})t_{L2}, \quad (14)$$

где r_{22} – это количество заданий в пакете ($i=2$)-го типа в ($j=2$)-й позиции в π^L , соответствующее составу ($h=1$)-го пакета из q_2 -го решения в матрице A_2^n . При разбиении одного из подмножеств G_{2q_2} на подмножества G_{3q_3} для каждого из них формируются вершины в дереве МВГ, которым соответствуют наборы $\langle 3, k_3, q_3, 1 \rangle$. При этом ($h=1$)-й ПЗ ($i=3$)-го типа из q_3 -го решения по составам ПЗ добавляется в π^l . Формируется матрица R соответствующего вида, на основе которой реализуется вычисление значений t_{ij}^{nl} ($i = \overline{1, 3}$, $j = \overline{1, 3}$, $l = \overline{1, L}$). Оценка $H_{3q_3}^H$ определяются следующим образом:

$$H_{3q_3}^H = t_{33}^{nL} + r_{33}t_{L,3} + t_{34}^L + n^4t_{L,4} + .. + t_{(n-1),n}^L + \\ + n^nt_{L,n} + t_{n,1}^L + (n^l - r_{11})t_{L1} + t_{1,2}^L + (n^2 - r_{22})t_{L2} + (n^3 - r_{33})t_{L3}, \quad (15)$$

где r_{33} – это количество заданий в пакете ($i=3$)-го типа в ($j=3$)-й позиции в π^L , соответствующее ($h=1$)-у пакету в q_3 -м решении в матрице A_3^n . С учетом (9), (10) выражение для вычисления нижних оценок $H_{iq_i}^h$ для множеств G_{iq_i} ($q_i = \overline{1, n_p^i}$) сформировано в виде:

$$H_{iq_i}^h = t_{ij}^{nL} + r_{ij} t_{L,i} + t_{i,i+1}^L + n^{i+1} t_{L,i+1} + \dots + t_{(n-1),n}^L + n^n t_{L,n} + t_{n,1}^L + (n^1 - r_{11}) t_{L1} + t_{1,2}^L + (n^2 - r_{22}) t_{L2} + \dots + (n^i - r_{ij}) t_{L,i}, \quad (16)$$

где r_{ij} – это количество заданий в пакете i -го типа в j -й позиции в π^L , соответствующее составу ($h=1$)-го пакета в q_i -м решении в A_i^n . Формирование матрицы R для вычисления значений $H_{iq_i}^h$ реализуется с использованием процедуры MATRICA_R1, рассмотренной ниже.

С целью обновления верхней оценки H^6 для подмножеств G_{iq_i} формируются допустимые решения по порядкам выполнения ПЗ в МС, предусматривающие, что в последовательностях π^l пакеты размещаются следующим образом (с учетом эвристики):

- первоначально в π^l ($l = \overline{1, L}$) размещаются ПЗ рассмотренных i' -х типов ($i' = \overline{1, i}$, i – тип заданий, для которого сформированы подмножества решений на текущей итерации алгоритма МВГ);
- в π^l ($l = \overline{1, L}$) размещаются фиксированные ПЗ оставшихся i' -х типов ($i' = \overline{i+1, n}$), для которых подмножества $G_{i'q_i}$ не сформированы;
- в π^l ($l = \overline{1, L}$) размещаются «оставшиеся» задания i' -х типов ($i' = \overline{1, i}$) (типов, для которых сформированы подмножества $G_{i'q_i}$).

Для построения матрицы R , соответствующей последовательностям π^l , сформированным с использованием предложенного способа, разработана процедура MATRICA_R2, рассмотренная ниже. С учетом

матрицы R , а также значений t_{ij}^{nl} ($i = \overline{1, n}$, $j = \overline{1, (2i + (n - i))}$, $l = \overline{1, L}$) вычисляются значения $f(\cdot)$, характеризующие подмножества G_{i, q_i} : $f(\cdot) = (t_{i, n_p}^{nL} + r_{i, n_p} t_{L, i})$ (при $i = \overline{1, n}$, $j = \overline{1, (2i + (n - i))}$). Проверяется условие $H^6 > \max f([M, A], [P, R, \{T^{nl} | l = \overline{1, L}\}])$, при выполнении которого значение H^6 (рекорда) переопределяется.

Предварительная инициализация входных параметров алгоритма МВГ оптимизации решения по составам ПЗ n типов имеет вид: задать идентификатор типа заданий i равным 1 ($i=1$); для ($i=1$)-типа заданий задать значение индекса q_1 строки в A_1^n ($q_1 = \overline{1, n_p^1}$), решение из которой будет интерпретироваться при формировании дерева; для ($i=1$)-типа заданий задать номер пакета h_1 равным 1. Алгоритм, реализующий процедуру МВГ с целью оптимизации решений по составам ПЗ разных типов, имеет следующий порядок шагов:

1. Для ($i=1$)-го типа заданий сформировать корневую вершину, сопоставить корневой вершине набор значений параметров $\langle i, k_s, q_i, h_i \rangle$ в виде: $\langle 1, 1, q_1, 1 \rangle$. Инициализировать $Q_1^1 : Q_1^1 = \{1\}$.

2. По формуле (8) вычислить значение верхней оценки H^6 значений критерия $f(\cdot)$ для рассматриваемого множества допустимых решений G_{1, q_1} (полученное значение H^6 – текущий рекорд).

3. Модифицировать тип заданий i и номер итерации алгоритма s (номер уровня дерева МВГ): $i=i+1$, $s=s+1$. Инициализировать H_s^6 значением W (W – бесконечно большое число): $H_s^6 = W$.

4. Если $i \leq n$, то инициализировать параметр h_i для рассматриваемого i -го типа заданий значением 1 ($h_i = 1$). Для рассматриваемого i -го типа заданий выполнить инициализацию параметра q_i значением 1 ($q_i = 1$). Задать значение счетчика sh_s количества потомков, закрепленных (зафиксированных) на s -ом уровне дерева МВГ, равным

0 ($sh_s = 0$). Инициализировать множество $Q_s^I = \emptyset$. Модифицировать множество $I_{MBГ}$: $I_{MBГ} = I_{MBГ} \cup \{i\}$. Перейти на пункт 5. Если $i > n$, то инициализировать H^6 : $H^6 = W$. Перейти на пункт 10.

5. Если $Q_{s-1}^I = \emptyset$, то проверить $H^6 > H_s^6$. Если $H^6 > H_s^6$, то инициализировать H^6 : $H^6 = H_s^6$. Перейти на пункт 3. Если $Q_{s-1}^I \neq \emptyset$, то извлечь из Q_{s-1}^I идентификатор родительской вершины k_{s-1}^* , для которой формируются потомки: $k_{s-1}^* = \min\{k_{s-1} / k_{s-1} \in Q_{s-1}^I\}$. Модифицировать множество Q_{s-1}^I : $Q_{s-1}^I = Q_{s-1}^I \setminus \{k_{s-1}^*\}$. Выполнить переход на $(s-1)$ -ом уровне дерева MBГ к вершине с номером k_{s-1}^* .

6. Модифицировать значение sh_s счетчика вершин, зафиксированных в дереве MBГ на s -ой итерации алгоритма: $sh_s = sh_s + 1$. Сформировать потомка вершины с идентификатором k_{s-1}^* . Поставить ей в соответствие набор параметров вида $\langle i, k_s, q_i, h_i \rangle$ (где $k_s = sh_s$).

7. Выполнить обход в глубину из корневой вершины в рассматриваемую k_s -ю вершину (в результате определяются значения параметров q_i для всех вершин, лежащих на пути ($i \in I_{MBГ}$)). Выполнить вызов процедуры MATRICA_R1 построения матрицы R для рассматриваемых типов заданий $i \in I_{MBГ}$ и значений q_i ($i \in I_{MBГ}$). Вычислить на основе матрицы R значение H_{i,q_i}^H для сформированного подмножества решений G_{i,q_i} (вершины k_s). Выполнить вызов процедуры MATRICA_R2 построения матрицы R для типов заданий $i \in I_{MBГ}$ и полученных значений q_i ($i \in I_{MBГ}$). Вычислить на основе матрицы R значение критерия $f([M, A], [P, R, \{T^{nl} / l = \overline{1, L}\}])$.

8. Если $H_{i,q_i}^H \geq H^6$, то удалить рассматриваемую вершину из дерева MBГ (в вершине с идентификатором k_{s-1}^* удалить ссылку на эту вершину). Модифицировать значение sh_s счетчика вершин, зафик-

сированных в дереве МВГ на s -й итерации алгоритма: $sh_s = sh_s - 1$. Выполнить проверку условия $q_i < n_p^i$. Если $q_i < n_p^i$, то модифицировать значение q_i : $q_i = q_i + 1$. Перейти на пункт 6. Если $q_i > n_p^i$, то положить $q_i = 1$. Перейти на пункт 5.

9. Если $H_{i,q_i}^h < H^e$, то модифицировать множество Q_s^l вершин, зафиксированных в дереве МВГ на s -ой итерации алгоритма: $Q_s^l = Q_s^l \cup \{k_s\}$. Если $H_s^e > f(\cdot)$, то изменить H_s^e : $H_s^e = f(\cdot)$. Если $H_s^e \leq f(\cdot)$, то проверить $q_i < n_p^i$. Если $q_i < n_p^i$, то модифицировать значение q_i : $q_i = q_i + 1$. Перейти на пункт 6. Если $q_i > n_p^i$, то положить $q_i = 1$. Перейти на пункт 5.

10. Если $Q_s^l \neq \emptyset$, то выделить в множестве Q_s^l , соответствующем последней итерации алгоритма построения дерева МВГ, вершину k_s^* : $k_s^* = \min\{k_s / k_s \in Q_s^l\}$. Модифицировать множество Q_s^l : $Q_s^l = Q_s^l \setminus \{k_s^*\}$. Если $Q_s^l = \emptyset$, то перейти на пункт 13.

11. Выполнить обход в глубину из корневой вершины в рассматриваемую k_s -ю вершину (определяются значения q_i для вершин, лежащих на пути). Вызвать процедуру MATRICA_R3 построения матрицы R для рассматриваемых типов заданий. Вычислить на основе матрицы R значение $f(\cdot)$ для рассматриваемого решения.

12. Если $f(\cdot) < H^e$, то зафиксировать параметры q_i вершин, соответствующих решениям по составам ПЗ i -х типов ($i = \overline{1, n}$): $q_i^h = q_i$ ($i = \overline{1, n}$). Модифицировать $H^e = f(\cdot)$. Перейти на шаг 10.

13. Останов алгоритма.

В результате получены индексы q_i^h ($i = \overline{1, n}$) решений по составам ПЗ i -х типов (номера q_i^h строк в матрицах A_i^n), комплексирование которых позволяет получить решение с минимальным значением $f(\cdot)$. После того, как для каждого дерева получено лучшее решение, в качестве оптимального выбирается решение с минимальным значением $f(\cdot)$. Вычислительная сложность алгоритма определена как

$O(n \cdot (n^i)^2)$. Для вычисления оценок H_{i,q_i}^n подмножеств G_{i,q_i} формируется матрица R с использованием процедуры MATRICA_R1. Входными параметрами для нее являются: множество $I_{MBГ}$; значения параметров q_i для вершин ($i \in I_{MBГ}$). Порядок ее шагов следующий:

1. Выполнить инициализацию множества I_1 типов заданий, используемого при реализации алгоритма: $I_1 = I_{MBГ}$.

2. Задать значения параметров h_i ($i \in I_1$) следующим образом: $h_i = 1$, инициализировать номер j позиции ПЗ в π^l значением 1 ($j=1$).

3. Определить в множестве I_1 тип заданий i' , параметры ПЗ которого используются при инициализации значения элемента $r_{i',j}$ матрицы R : $i' = \min(i | i \in I_1)$, модифицировать I_1 : $I_1 = I_1 \setminus \{i'\}$.

4. Инициализировать значение элемента $r_{i',j}$ матрицы R : $r_{i',j} = a_{q_i, h_i}$. Модифицировать номер j позиции ПЗ в π^l : $j=j+1$.

5. Если $I_1 \neq \emptyset$, то перейти на пункт 3. Если $I_1 = \emptyset$, то пункт 5.

6. Останов алгоритма.

Процедура MATRICA_R2 реализует построение матрицы R , используемой при вычислении значений $f(\cdot)$ на s -ой итерации МВГ для допустимых решений. Ее входные параметры: множество $I_{MBГ}$; значения q_i для вершин, соответствующих решениям по составам ПЗ i -х типов ($i \in I_{MBГ}$); множество I типов заданий. Порядок шагов алгоритма процедуры MATRICA_R2 следующий:

1. Выполнить инициализацию: множества I : $I = \{1, 2, \dots, n\}$, множества I_1 : $I_1 = I_{MBГ}$, множества I_2 : $I_2 = I \setminus I_{MBГ}$.

2. Задать значения параметров h_i ($i \in I_1$) следующим образом: $h_i = 1$, инициализировать номер позиции j ПЗ в π^l значением 1 ($j=1$).

3. Определить в множестве I_1 тип заданий i' , параметры ПЗ которого используются при инициализации значения элемента $r_{i',j}$ матрицы R : $i' = \min(i | i \in I_1)$, модифицировать I_1 : $I_1 = I_1 \setminus \{i'\}$.

4. Инициализировать значение элемента $r_{i',j}$ матрицы R : $r_{i',j} = a_{q_i, h_i}$. Модифицировать номер j позиции ПЗ в π^l : $j=j+1$.

5. Если $I_1 = \emptyset$, то перейти на пункт 5. Если $I_1 \neq \emptyset$, то перейти на пункт 3.

6. Если $I_2 \neq \emptyset$, то извлечь из I_2 тип заданий, фиксированный пакет которых размещается в $\pi^l: i' = \min(i/i \in I_2)$. Инициализировать значение элемента $r_{i'j}$ матрицы $R: r_{i'j} = n^i$. Перейти на пункт 6. Если $I_2 = \emptyset$, то модифицировать номер j позиции: $j=j+1$. Инициализировать множество $I_1: I_1 = I_{MBГ}$. Перейти на пункт 7.

7. Модифицировать номер $j: j=j+1$. Перейти на пункт 5.

8. Определить в I_1 тип заданий i' , размещаемых в $\pi^l: i' = \min(i/i \in I_1)$, модифицировать $I_1: I_1 = I_1 \setminus \{i'\}$.

9. Инициализировать значение элемента $r_{i'j}$ матрицы $R: r_{i'j} = n^i - \sum_{s=1}^{j-1} r_{i',s}$. Модифицировать номер j позиции ПЗ в $\pi^l: j=j+1$.

10. Если $I_1 = \emptyset$, то перейти на пункт 10. Если $I_1 \neq \emptyset$, то перейти на пункт 7.

11. Останов алгоритма.

Процедура MATRICA_R3 реализует построение матрицы R для получения окончательного значения критерия $f(\cdot)$ для решений, получаемых для отдельных деревьев МВГ. Перед реализацией процедуры множества I_1 и I_2 заданы следующим образом: $I_1 = I_2 = I$, где I – упорядоченное множество типов заданий. Порядок шагов алгоритма:

1. Задать значения параметров $h_i (i = \overline{1, n})$: $h_i = 1$ при $(i = \overline{1, n})$, инициализировать номер позиции j ПЗ в π^l значением 1 ($j=1$).

2. Определить в множестве I_2 тип заданий i' , параметры ПЗ которого используются при инициализации значения элемента r_{ij} матрицы $R: i' = \min(i/i \in I_2)$, модифицировать $I_2: I_2 = I_2 \setminus \{i'\}$.

3. Инициализировать значение элемента $r_{i'j}$ матрицы $R:$

$$r_{i'j} = a_{q_{i'}, h_{i'}}, \text{ Если } \sum_{h=1}^{n_p} r_{i'h} < n^i, \text{ то } h_{i'}: h_{i'} = h_{i'} + 1, \text{ модифицировать } j:$$

$j=j+1$. Перейти на пункт 4. Если $\sum_{h=1}^{n_p} r_{i'h} = n^i$, то $I_1 = I_1 \setminus \{i'\}$, модифицировать $j: j=j+1$. Перейти на пункт 4.

4. Если $I_2 \neq \emptyset$, то определить тип заданий i' , пакет которых размещается в $\pi^l: i' = \min(i/i \in I_2)$. Перейти на пункт 3. Если $I_2 = \emptyset$, то проверить условие $I_1 \neq \emptyset$. Если $I_1 \neq \emptyset$, то инициализировать $I_2: I_2 = I_1$. Перейти на пункт 2. Если $I_1 = \emptyset$, то перейти на пункт 5.

5. Останов алгоритма.

Для исследования эффективности планирования с использованием рассмотренной реализации МВГ использованы параметры: неоднородность длительностей выполнения заданий на приборах МС – $\max(t_{ij}) / \min(t_{ij})$; неоднородность длительностей переналадок приборов МС – $\max(t_{ij}^l) / \min(t_{ij}^l)$. Значения $\max(t_{ij}^l) / \min(t_{ij}^l)$ задавались равными 1, 2, 4, 8, 16; значения $\max(t_{ij}) / \min(t_{ij})$ – равными 1, 2, 4, 8, 16. Значения параметра n заданы равными 3 и 5, параметра n^i равными 6,8,12. Значение L задано равным 5. В качестве параметра, характеризующего снижение времени на выполнение ПЗ, использовалось отношение $f_{эмосп} = (f^{\text{фикс}} - f^{\text{мосп}}) / f^{\text{фикс}}$, где $f^{\text{фикс}}$ – значение критерия f для фиксированных пакетов (ФП), $f^{\text{мосп}}$ – значение критерия f для решения с оптимизированными составами ПЗ. Графики, отображающие зависимости снижения времени выполнения заданий при формировании ПЗ с использованием МВГ по сравнению с фиксированными пакетами для $n=5$, $n^i \in \{6,8,12\}$, представлены на рисунках 1-3. Аналогичные результаты получены при исследовании зависимости $f_{эмосп}$ от входных параметров при $n=5$, $n^i \in \{6,8,12\}$.

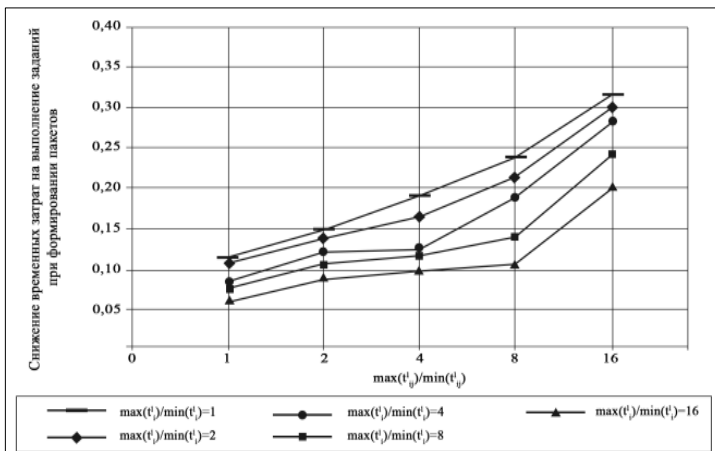


Рис. 1. Снижение времени выполнения заданий в МС при оптимизации составов ПЗ по сравнению с ФП ($n = 5, n^i = 6$)

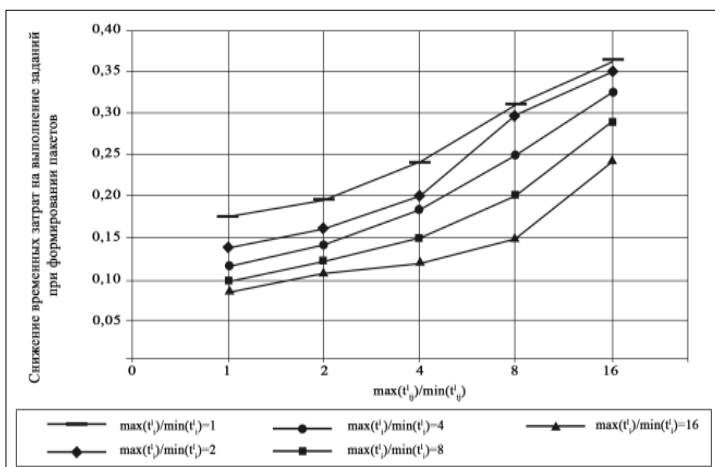


Рис. 2. Снижение времени выполнения заданий в МС при оптимизации составов пакетов по сравнению с ФП ($n = 5, n^i = 8$)

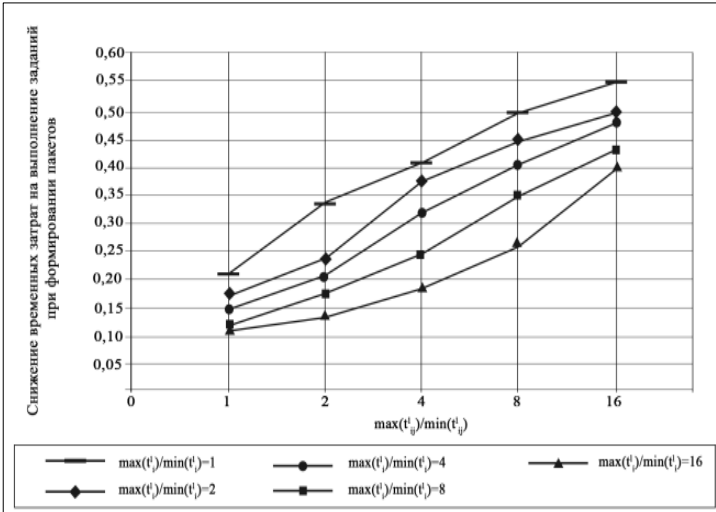


Рис. 3. Снижение времени выполнения заданий в МС при оптимизации составов пакетов по сравнению с ФП ($n = 5, n^i = 12$)

Результаты исследований показали, что использование МВГ при планировании позволяет максимально на 45% (в среднем на 30%) снизить временные затраты на выполнение ПЗ в МС (в сравнении с фиксированными ПЗ).

4. Заключение. Существующие методы планирования выполнения ПЗ в МС позволяют определять решения задач по составам ПЗ и расписаниям их выполнения малой размерности, не гарантируют получение решений, приближающихся к оптимальным. В то же время с использованием МВГ реализуется построения расписаний выполнения ЕЗ в системах разного вида. Для использования МВГ с целью оптимизации составов ПЗ разработано эвристическое правило упорядочивания пакетов в последовательностях их выполнения в МС. Разработана математическая модель процесса выполнения ПЗ в МС. Построены выражения для вычисления нижних и верхних оценок значений критерия для каждого из подмножеств решений, сформированных в результате реализации процедуры разбиения МВГ. Сформулирован алгоритм формирования решений по составам ПЗ разных типов и алгоритм МВГ, позволяющий определить оптимальное решение по составам ПЗ. Программная реализация алгоритмов показала, что их использование позволяет увеличить эффективность планирования в среднем на 30% по сравнению с ФП. Результатами, представленными в работе и обла-

дающими элементами научной новизны, являются: математическая модель процесса выполнения ПЗ в МС; эвристическое правило упорядочивания ПЗ в последовательностях их выполнения в МС; алгоритм МВГ, адаптированный для решения задачи оптимизации составов ПЗ, выполняемых в МС; выражение для расчета нижних оценок и верхней оценки; способ построения последовательностей выполнения ПЗ в МС.

Литература

1. Ogun B., Cigdem A.-U. Mathematical Models for a Batch Scheduling Problem to Minimize Earliness and Tardiness. *Journal of Industrial Engineering and Management*. JIEM, 2018. № 11(3). pp. 390-405.
2. Chaudhry I.A., Elbadawi I. A-Q., Usman M., Chughtai M.T. Minimising Total Flowtime in a No-Wait Flow Shop (NWFS) using Genetic Algorithms. *Ingeniería e Investigación*. 2018. Vol. 38. № 3. pp. 68-79.
3. Tan Y., Huang W., Sun Y., Yue Y. Comparative Study of Different Approaches to Solve Batch Process Scheduling and Optimisation Problems. *Proceedings of the 18th International Conference on Automation & Computing*. Loughborough University, Leicestershire. UK. 2012. pp 424-444.
4. Кротов К.В. Использование аппарата генетических алгоритмов при формировании решений по составам партий данных в двухуровневой задаче построения комплексных расписаний их обработки. *Автоматизированные технологии и производства // Международный научно-технический журнал*. 2017. №2 (16). С. 23-34.
5. Li X.L., Wang Y. Scheduling Batch Processing Machine Using Max–Min Ant System Algorithm Improved by a Local Search Method. *Mathematical Problems in Engineering*. 2018. Vol. 2018.
6. Li Sh., Cheng T.C.E., Ng C.T., Yuan J. Single-machine batch scheduling with job processing time compatibility. *Theoretical Computer Science*. 2015. Vol. 583. pp. 57-66.
7. Jin M., Liu X., Luo W. Single-Machine Parallel-Batch Scheduling with Nonidentical Job Sizes and Rejection. *Mathematics*. 2020. Vol. 8.
8. Surjandari I., Rachman A., Purdianta, Dhini A. The batch scheduling model for dynamic multi-item, multi-level production in an assembly job shop with parallel machines. *International Journal of Technology*. 2015. № 1. pp. 84-96.
9. Joglekar G. Using Simulation for Scheduling and Rescheduling of Batch Processes. *Processes*. 2017. № 5.
10. Ковалев М.Я. Модели и методы календарного планирования. Курс лекций. Минск: БГУ. 2004. 63 с.
11. Morrison D.R., Jacobson Sh.H., Sauppe J.J., Sewell E.C. Branch-and-bound algorithms: A survey of recent advances in searching, branching and pruning. *Discrete Optimization*. 2016. № 19. pp. 79-102.
12. Dawd S.T., Ayvaz B. A branch and bound approach for single machine scheduling problem. *Istanbul Commerce University. Journal of Science*. 2017. № 16 (31). pp. 43-55.
13. Rasti-Barzoki M., Hejazi S.R. A branch and bound algorithm to minimize the total weighted number of tardy jobs and delivery costs with late deliveries for a supply chain-scheduling problem. *Journal of Industrial and Systems Engineering*. 2017. Vol. 10. № 1. pp 50- 60.

14. Прилуцкий М.Х., Власов В.С. Метод ветвей и границ с эвристическими оценками для конвейерной задачи теории расписаний // Математическое моделирование. Оптимальное управление. Вестник Нижегородского университета им. Н.И. Лобачевского. 2008. № 3. 147-153 с.
15. Takano M.I., Nagano M.S. A branch-and bound method to minimize the makespan in a permutation flow shop with blocking and setup times. *Cogent Engineering*. 2017.
16. Григорьева Н.С. Алгоритм ветвей и границ для задачи составления расписания на параллельных процессорах // Вестник Санкт-Петербургского университета. серия 10. 2009. выпуск 1. 44-55 с.
17. Mazda Ch.N., Kurniawati D.A. Branch and Bound Method to Overcome Delay Delivery Order in Flow Shop Scheduling Problem. *IOP Conf. Series: Materials Science and Engineering*. 2020.
18. Watermeyer K., Zimmermann J. A branch-and-bound procedure for the resource-constrained project scheduling problem with partially renewable resources and general temporal constraints. *OR Spectrum*. 2020. № 42. pp. 427-460.
19. Hu Sh., Wang S., Kao Y., Ito T., Sun X. A Branch and Bound Algorithm for Project Scheduling Problem with Spatial Resource Constraints. *Hindawi Publishing Corporation. Mathematical Problems in Engineering*. Vol. 2015. P. 9.
20. Могилев А.А. Обзор методов решения задач теории расписаний // Информатика, вычислительная техника и инженерное образование. 2019. № 4 (37). 19-32 с.
21. Кротов К.В. Информационная модель многоуровневой системы выполнения конвейеризированных программ // Международный научно-технический журнал «Проблемы управления и информатики». 2014. № 3. 89-101 с.
22. Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G., Shmoys D.B. Sequencing and Scheduling: Algorithms and Complexity. *Handbook in Operations Research and Management Science*. North-Holland, Amsterdam. 1993. Vol. 4. pp. 445-522.
23. Кротов К.В. Комплексный метод определения эффективных решений по составам партий данных и расписаниям их обработки в конвейерных системах // Журнал «Вычислительные технологии». Новосибирск. Изд-во Института вычислительных технологий СО РАН. 2018. № 3. 58-76 с.
24. Кротов К.В., Скатков А.В. Организация web-ориентированного сервиса мониторинга окружающей среды с использованием данных дистанционного зондирования Земли и конвейеризации обработки данных // Труды учебных заведений связи. 2021. Т. 7. №1. 105-121 с.

Кротов Кирилл Викторович — канд. техн. наук, доцент, Кафедра информационных систем ФГАОУ ВО Севастопольский Государственный университет (СевГУ). Область научных интересов: моделирование процессов в вычислительных системах, оптимальное планирование процессов в вычислительных системах, иерархические игры. Число научных публикаций — 110. krotov_k1@mail.ru; ул. Университетская, 33, 299053, Севастополь, Россия; р.т.: +7 (978) 730-38-19.

K. KROTOV

MATHEMATICAL MODEL AND ALGORITHM OF BRANCH AND BOUNDARY METHOD FOR OPTIMIZING SOLUTIONS FOR PACKAGE COMPOSITIONS IN MULTI-STAGE SYSTEMS

Krotov K.V. Mathematical model and algorithm of branch and boundary method for optimizing solutions for package compositions in multi-stage systems

Abstract. Modern methods for solving problems of planning of task packages execution in multi-stage systems are characterized by the presence of restrictions on their dimension, the impossibility of obtaining guaranteed best results in comparison with fixed packages for different values of the input parameters of tasks. The problem of optimizing the composition of task packages executed in multi-stage systems using the method of branches and borders is solved in the paper. Studies of various ways of forming the order of execution of task packages in multi-stage systems (heuristic rules for ordering task packages in the sequences of their execution on MS devices) have been carried out. The method of ordering packets in the sequence of their execution (a heuristic rule), which minimizes the total time for implementing actions with them on the devices, is defined. The method of ordering the types of tasks, according to which their packages are considered in the procedure of the method of branches and borders, is formulated on the basis of the obtained rule. A mathematical model of the process of implementing actions with packages on the system devices, which provides the calculation of its parameters, has been built. The construction of a method for forming all possible solutions for the composition of task packages for a given number of them has been completed. Decisions on the composition of task packages of different types are interpreted in the procedure of the method of branches and borders in order to build the optimal combination of them. To implement the method of branches and borders, a branching (splitting) procedure is formulated, which assumes the formation of subsets of solutions that include packages of different compositions of tasks of the same type. Expressions for calculating the lower and upper estimates of the values of the optimization criterion for the composition of packages for subsets formed in the branching procedure are constructed. The dropout procedure involves the exclusion of subsets whose lower estimate is not less than the record. To find optimal solutions, a breadth-first search strategy is applied, which provides for the study of all subsets of solutions that include various packages of tasks of the same type obtained as a result of the procedure for splitting subsets of tasks that are not excluded from consideration after the implementation of the dropout procedure. The developed algorithms are implemented programmatically, which allowed to obtain the results of planning the execution of task packages in a multi-stage system, which are on average 30 % better than fixed packages.

Keywords: multi-stage system, task packages, method of branches and boundaries, heuristic rule, schedules.

Krotov Kirill — Ph.D., Assistant Professor, Information systems department Sevastopol State University. Research interests: modeling of processes in computing systems, optimal planning of processes in computing systems, hierarchical games. The number of publications — 110. krotov_k1@mail.ru; 33, St. University, 299053, Sevastopol, Russia; office phone: +7 (978) 730-38-19.

References

1. OgunB., Cigdem A.-U. Mathematical Models for a Batch Scheduling Problem to Minimize Earliness and Tardiness. *Journal of Industrial Engineering and Management*. JIEM, 2018. № 11(3). pp. 390-405.
2. Chaudhry I.A., Elbadawi I. A.-Q., Usman M., Chughtai M.T. Minimising Total Flowtime in a No-Wait Flow Shop (NWFS) using Genetic Algorithms. *Ingeniería e Investigación*. 2018. Vol. 38. № 3. pp. 68-79.
3. Tan Y., Huangi W., Sun Y., Yue Y. Comparative Study of Different Approaches to Solve Batch Process Sheduling and Optimisation Problems. *Proceedings of the 18th International Conference on Automation & Computing*. Loughborough University. Leicestershire. UK. 2012. pp 424–444.
4. Krotov K.V. [The use of the apparatus of genetic algorithms in the formation of decisions on the composition of data batches in the two-level task of constructing complex schedules for their processing]. *Avtomatizirovannye tekhnologii i proizvodstva. Mezhdunarodnyj nauchno-tehnicheskij zhurnal. – Automated technologies and production. International scientific and technical journal*. 2017. № 2 (16), pp. 23-34 (In Russ).
5. Li X.L., Wang Y. Scheduling Batch Processing Machine Using Max–Min Ant System Algorithm Improved by a Local Search Method. *Mathematical Problems in Engineering*. 2018. Vol. 2018.
6. Li Sh., Cheng T.C.E., Ng C.T., Yuan J. Single-machine batch scheduling with job processing time compatibility. *Theoretical Computer Science*. 2015. Vol. 583. pp. 57-66.
7. Jin M., Liu X., Luo W. Single-Machine Parallel-Batch Scheduling with Nonidentical Job Sizes and Rejection. *Mathematics*. 2020. Vol. 8.
8. Surjandari I., Rachman A., Purdianta, Dhini A. The batch scheduling model for dynamic multi-item, multi-level production in an assembly job shop with parallel machines. *International Journal of Technology*. 2015. № 1. pp. 84-96.
9. Joglekar G. Using Simulation for Scheduling and Rescheduling of Batch Processes. *Processes*. 2017. № 5.
10. Kovalev M.M. [Models and methods of calendar planning: a course of lectures]. Minsk: Izdatel'stvo BGU. 2004. P. 63. (In Russ).
11. Morrison D.R., Jacobson Sh.H., Sauppe J.J., Sewell E.C. Branch-and-bound algorithms: A survey of recent advances in searching, branching and pruning. *Discrete Optimization*. 2016. № 19. pp. 79-102.
12. Dawd S.T., Ayvaz B. A branch and bound approach for single machine scheduling problem. *Istanbul Commerce University. Journal of Science*. 2017. № 16 (31). pp. 43-55.
13. Rasti-Barzoki M., Hejazi S.R. A branch and bound algorithm to minimize the total weighted number of tardy jobs and delivery costs with late deliveries for a supply chain-scheduling problem. *Journal of Industrial and Systems Engineering*. 2017. Vol. 10. № 1. pp 50- 60.
14. Prilutsky M.H., Vlasov V.S. [The method of branches and boundaries with heuristic estimates for the pipeline problem of the theory of schedules.] *Matematicheskoe modelirovanie. Optimal'noe upravlenie. Vestnik Nizhegorodskogo universiteta im. N.I. Lobachevskogo. [Mathematical modeling. Optimal control. Bulletin of the Nizhny Novgorod University named after N. I. Lobachevsky]*. 2008. № 3. pp. 147-153 (In Russ).
15. Takano M.I., Nagano M.S. A branch-and bound method to minimize the makespan in a permutation flow shop with blocking and setup times. *Cogent Engineering*, 2017. 4:1. 1389638. DOI: 10.1080/23311916.2017.1389638.

16. Grigorieva N.S. [Algorithm of branches and boundaries for the task of scheduling on parallel processors]. Vestnik Sankt-Peterburgskogo universiteta, seriya 10.–Vestnik of St. Petersburg University, series 10, 2009, issue 1, pp. 44-55(In Russ).
17. Mazda Ch.N., Kurniawati D.A. Branch and Bound Method to Overcome Delay Delivery Order in Flow Shop Scheduling Problem. IOP Conf. Series: Materials Science and Engineering, 2020. IOP Publishing, DOI:10.1088/1757-899X/1003/1/012129.
18. Watermeyer K., Zimmermann J. A branch-and-bound procedure for the resource-constrained project scheduling problem with partially renewable resources and general temporal constraints. OR Spectrum, 2020. № 42. p. 427–460, DOI: 10.1007/s00291-020-00583-z.
19. Hu Sh., Wang S., Kao Y., Ito T., Sun X. A Branch and Bound Algorithm for Project Scheduling Problem with Spatial Resource Constraints. Hindawi Publishing Corporation, Mathematical Problems in Engineering, Volume 2015. Article ID 628259. 9 pages. DOI: 10.1155/2015/628259.
20. Mogilev A.A. [Review of methods for solving problems in the theory of schedules]. Informatika, vychislitel'naya tekhnika i inzhenerное obrazovanie. [Informatics, Computer Engineering and Engineering education]. 2019. №. 4 (37), pp. 19-32 (In Russ).
21. Кротов К.В. Информационная модель многоуровневой системы выполнения конвейеризированных программ // Международный научно-технический журнал «Проблемы управления и информатики». 2014. № 3. 89-101 с.
22. Lawler E.L., Lenstra J.K., Rinnooy Kan A.H.G., Shmoys D.B. Sequencing and Scheduling: Algorithms and Complexity. Handbook in Operations Research and Management Science. North-Holland, Amsterdam, 1993. Vol. 4. pp. 445-522.
23. Krotov K.V. [A complex method for determining effective solutions for the composition of data batches and schedules of their processing in conveyor systems]. Zhurnal «Vychislitel'nye tekhnologii», Izd-vo Instituta vychislitel'nyh tekhnologij SO RAN. [Computational technologies, Publishing house Of the Institute of computational technologies SB RAS]. Vol. 23. № 3. 2018. pp. 58-76 (In Russ).
24. Krotov K.V., Skatkov A.V. [Organization of a web-based environmental monitoring service using Earth remote sensing data and data processing pipelining]. Trudy uchebnyh zavedenij svyazi. [Proceedings of educational institutions of Communications]. 2021. Vol. 7. No. 1. pp. 105-121 (In Russ).

V. ZHMUD, L. DIMITROV, G. SABLINA, H. ROTH, J. NOSEK, W. HARDT
**ON THE EXPEDIENCY AND POSSIBILITIES OF
APPROXIMATING A PURE DELAY LINK**

Zhmud V., Dimitrov L., Sablina G., Roth H., Nosek J., Hardt W. On the Expediency and Possibilities of Approximating a Pure Delay Link.

Abstract. When solving problems of controlling an object with delay, it is often necessary to approximate a pure delay link with a minimum phase link in order to ensure the possibility of using analytical methods for regulator design. There are many approximation methods based on the Taylor series expansion, as well as modified methods. The most famous one is the Padé approximation method. The known approximation methods have significant drawbacks, which this paper reveals. However, there are other methods of forming other types of filters that can serve as a better approximation in determining the delay relationship, although they are not used for these purposes. In particular, methods of forming the desired differential equation of a locked-loop system of a given order by the method of numerical optimization are known. In this case, the locked-loop system behaves like a filter of the corresponding order, the numerator of which is equal to one, and the specified polynomial is in the denominator. Modeling has shown that such a filter is an effective alternative approximation of the delay link and can be used for the same purposes for which it was supposed to use the Padé approximation. The polynomial coefficients in the literature were calculated only up to the 12th order. The higher the polynomial order is, the more accurate the approximation is.

Keywords: Padé formula, delay, approximation, control, automation.

1. Introduction. The development of numerical methods makes it possible to easily and efficiently design a controller for a locked-loop system if the mathematical model of the object is known with sufficient accuracy (as a rule, it is sufficient to know the basic parameters of the model with an error of at least 1%). However, many research teams are actively developing analytical methods for the synthesis of controllers, including for objects containing a pure delay link in their model. This does not seem to be extremely relevant since such software tools as MATLAB, Simulink, MathCAD, SimInTech, VisSim allow simulation of a pure delay link without any approximations. Nevertheless, papers on such an approximation are published [1–7], dissertations are defended [8–10], and this direction is widely lobbied by various research teams [11–17]. Considerations against using such an approximation are as follows: a) a pure delay link is characterized by a linear dependence of the delay on frequency, any approximation by a filter cannot provide such a dependence; b) well-known software tools and methods for designing regulators based on their use do not require any approximation, as they easily simulate the delay link inaccuracy; c) each approximation introduces an error in the calculations and the results of these calculations. Therefore, the design result with such an approximation may contain such a significant error that it will be inapplicable.

However, it would be wrong to reject any research carried out using such an approximation since a simple simulation experiment shows that in a certain area of problems this approximation allows obtaining the required result in a fairly simple way. In this case, it is necessary not to reject any such approximation completely, but it would be useful to determine the limits of its applicability for solving problems of controller design, and this is the purpose of this paper. The approximation of a pure delay link by a transfer function in the form of a rational fraction is still widely used to implement the possibilities of designing controllers for objects with delay [1–17]. Almost all real objects have a delay, and analytical design methods are still used by many researchers (despite the rapid development of numerical methods [18–22]), the problem of such an approximation is of interest at least from the standpoint of an adequate assessment of the results, usefulness and scientific contribution of dissertations and new publications using such an approximation. The Padé approximation is often called the most adequate in the literature, arguing that it best meets the design objectives of the controller [11–17]. This statement is justified for objects with large inertia that is not associated with this delay. But this is not always the case. There are object models for which such an approximation cannot be used. For example, if the object model is in a series connection of an integrator and a delay link, then the Padé approximation is not efficient enough. This could be verified by numerical modeling. In the case of application of the method of numerical optimization, such an approximation does not allow solving the problem of regulator design for such objects.

2. Task Statement. Let us consider a control object whose mathematical model has a dependence of the output value $y(t)$ on the input action $u(t)$, known in the form of the transfer function $W(s)$, which in addition to the minimum phase part also contains a link of pure delay $e^{-\tau s}$. Here s is the argument of the Laplace transform. The object is affected by an unknown disturbance $h(t)$, which causes a change in the output value $\Delta y(t)$. The effect of this disturbance can be described as an uncontrolled addition to the output value, which gives the output a modified value of $z(t)$. It is necessary to design a control system for this object so that the output value repeats the prescription $v(t)$ as accurately as possible and, as far as possible, does not depend on interference $h(t)$. This is the classical formulation of the automatic control problem, which is most successfully solved by creating a loop with a negative feedback unit with a controller in a front loop, as shown in Figure 1. As a rule, the problem of design involves finding a mathematical model of the controller. If the controller is sought in the form of a PID controller, i.e. a structure containing proportional, integrating and derivation links connected in parallel so that the output signals are summed

and fed to the object, then the design problem comes down to finding the coefficients of these links. As a rule, the controller should provide the highest achievable speed, zero static error and, if possible, the minimum overshoot, no more than 10%.

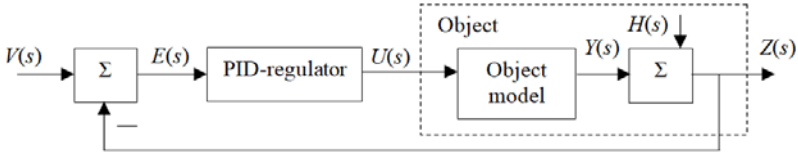


Fig. 1. A typical automatic control system with a sequential PID controller in a direct loop: values are given in terms of Laplace transforms of time functions

Among the methods for solving this problem can be distinguished analytical and numerical methods. Analytical methods deal with differential equations. The presence of a delay link complicates the application of such methods, as does the presence of nonlinearities and other problematic elements. Numerical modeling methods are effective for any of the most complex models of an object. In this case, the presence of any nonlinearities delays or other features of the object model is not in the list of problems, as all these elements are easily simulated by any of the above software.

The paper [21] provides information on two types of approximation. The first one is based on the expansion of the transfer function of the pure delay link in the Taylor series. The second one is the Padé approximation which differs in the presence in the numerator of the transfer function of a polynomial of the same order as in the denominator, but the terms with an odd power of the argument, Laplace transforms in this model have a negative sign. Paper [21] proposed its own approximation, which is said to be free from the drawbacks of the two indicated models. Throughout the entire cycle of works, including [1–22], the approximation is used to calculate the controller for the object with transfer function containing the following multiplier:

$$W_D(s) = e^{-\tau s}. \quad (1)$$

Here τ is the time constant. Various models of the transfer function have been proposed, and according to many authors, this one is the closest one to the transfer function of this link. The transfer function (1) is generally substituted by a rational fraction in the form of the ratio of two polynomials in the argument s .

$$W_M(s) = \frac{B(s)}{D(s)} = \frac{\sum_{i=0}^m b_i s^i}{\sum_{j=0}^l a_j s^j}. \quad (2)$$

Here $B(s)$ and $D(s)$ are polynomials in s . The degrees of these polynomials are equal to m and l respectively. For the transfer function (1) to be physically realizable, the order of the numerator must be less than the order of the denominator: $l > m$. However, for controllers and other elements that have much faster performance than other elements in the same control system, it is allowed that the order of the numerator is equal to the order of the denominator: $l = m$. In some cases, transfer functions are used in which the order of the denominator is one unit larger than the order of the numerator. For example, in an ideal derivative link: $l=m+1$.

Taylor's method is based on power series. For the transfer function of the delay link (1), the Taylor series expansion for a function of the form (1) gives:

$$W_T(s) = \frac{1}{1 + \tau s + \frac{1}{2}(\tau s)^2 + \frac{1}{6}(\tau s)^3 + \dots}. \quad (3)$$

Many publications, including [12, 14, and 21], preference is given to the Padé approximation, which has the following form:

$$W_t^N(s) = \frac{\sum_{k=0}^N \frac{(n+k)!}{k!(n-k)!} (-s\tau)^{n-k}}{\sum_{k=0}^N \frac{(n+k)!}{k!(n-k)!} (s\tau)^{n-k}}. \quad (4)$$

Here N is the order of approximation of the model, n is the order of the polynomial in the denominator of the model, in this case, $N = n$.

The exact value of the second and subsequent coefficients in (4) depends on the order of approximation. The reason for this preference is not always clear. In [21] it is mentioned that the possibilities of the second option are greater, as it has a larger number of variable coefficients. We can't agree with this because the number of varying coefficients in relations (3) and (4) for the same order of the denominator exactly coincides, because in relation (4) the coefficients in the numerator in absolute value coincide with the coefficients of the denominator, and their sign alternates, starting with a positive one. If, in this case, the denominator of relation (4) is specified, then its numerator is also completely finally specified and no additional

variation of the coefficients is possible. Paper [21] suggests the model in the following form:

$$W_{VIM}(s) = \frac{b_0 + b_1 s}{1 + a_1 s + a_2 s^2}. \quad (5)$$

The unexpected fact is that in model (5) the free term of the numerator is given by a letter and it is determined in the denominator by a specific value equal to one. The fact is that the ratio of free terms determines the steady-state value of the response, and for relation (1) it is strictly equal to one. Therefore, the free term in the numerator must also be equal to one, and it does not need to be calculated by any method, $b_0=1$. It is obvious that for the case $\tau=1$ the paper [21] has obtained a relation where this is fulfilled:

$$W_{VIM}(s) = \frac{1 - 0.317s}{1 + 0.683s + 0.184s^2} \quad (6)$$

Figure 2 shows the transient processes when a unit step is applied to elements that have transfer functions (1), (3), (4) and (6) for $N = 4$. The response of the link with the transfer function (1) is ideal. The response according to relation (3) is excessively stretched in time. The response according to relation (4) has a very large starting value (which cannot take place in the original model) and a very large reverse overshoot. In this case, the response of link (6) seems to be most closely similar to the response of an ideal link (1). However, it is far from the ideal case, it stands at an intermediate position between the responses of elements (3) and (4), combining their disadvantages and advantages equally. The undoubted advantage of this model is the fact that at the zero moment the response is zero, which is more consistent with the true model (1).

This task aims to address the following issues:

- A. How appropriate is the use of approximation?
- B. Is not it better to abandon the use of methods that cannot do it without approximation, and to use only those methods that use an exact delay model?
- C. If the proposed approximation methods can still be used, then under what conditions?
- D. If it is appropriate to use the approximation, is it possible to improve it to expand the scope of its applicability, and if so, how exactly?

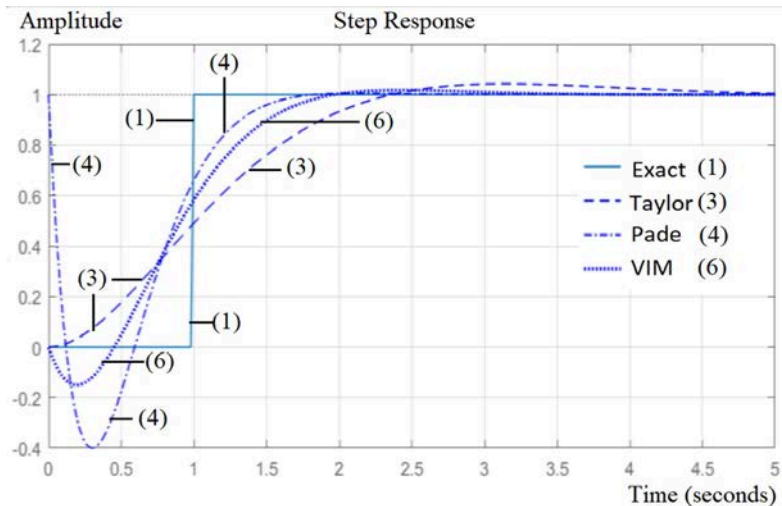


Fig. 2. Comparison of responses of models (1), (3), (4) and (6) from publication [21], order of models $N = 2$, the designation number is the same as the equation number

3. The Proposed Research Method. It is proposed to solve the problem by a method of numerical optimization. The graphical programming structure for solving this problem is shown in Figure 3. It contains a system model, a setting signal generator, a cost function estimator, an optimizer, and an oscilloscope. In this structure, an exact model of the object is used, and no approximations are required. To address the question of how competent the delay approximation is when using analytical methods, it is proposed to use this approximation when using the method of numerical optimization and then simulate the system with the obtained controller and use the exact model of the object. If the result of such a simulation is sufficiently similar to the result of the calculation by the model, it can be argued that such an approximation can be used. But if the result obtained by the numerical optimization method is still better than the result obtained using the approximation, we can supplement this conclusion with the statement that although the approximation can be used, it still does not give an optimal solution; thus, it is inappropriate.

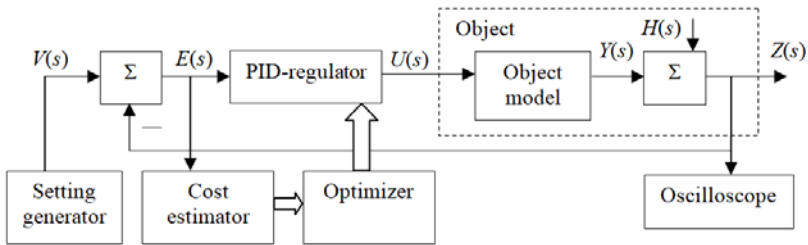


Fig. 3. The structure of graphical programming for solving the problem of numerical optimization of the controller

For modeling, the VisSim software was chosen for the reason that the calculation method in it coincides with the method of operation of any digital controller. Namely, the simulation is carried out in steps. This simplest advantage turns out to be decisive in comparison, for example, with the use of the MATLAB software, which allows performing analytical calculations in the form of finding the necessary functions, but at the same time, it does not form those delays that occur in a real system due to the real operation of a digital controller, and it does not form those errors in the estimation of the derivative and integral of the function by its readings, which also occur in a real system with a digital controller. Many problems can be specified that the MATLAB program solves efficiently and accurately, but the result obtained is not applicable in practice since such accuracy is not achieved in a real system. With the use of the VisSim software, this does not happen: if the software gives a result, the same result will take place in practice since in a real system the same step-by-step algorithms for calculating errors and control signals will be used.

4. Results of Studies. Comparison of the type of transient processes, which are the model's response to a single stepwise jump, shows that the models (3), (4) and (6) are very far from ideal. Indeed, the response of the Taylor model (3) is unnecessarily prolonged in time (see Figure 1), the response of the Padé model (4) is characterized by two drawbacks: firstly, a nonzero value at the moment $t = 0$, equal to the amplitude of the input signal, and secondly, significant reverse overshoot, reaching 40% of this amplitude. The response of the model (6) seems to be the best at first glance, as it starts from zero, the reverse overshoot is less than in the response of the model (4) and amounts to 15% of the input signal amplitude. In terms of rising time, this response occupies an intermediate position between the responses of the models (3) and (4). However, this characteristic is insufficient for making a decision that the model (6) is the best. Filters are described in [22]:

$$W_{CH}(s) = \frac{1}{D_{CH}(s)}. \quad (7)$$

The denominator of the fraction (7) contains polynomials. The lowest and highest coefficients are strictly equal to one. The remaining coefficients are found by numerical optimization using the following objective function:

$$F(T) = \int_0^T \left\{ k_w f \left[e(t) \frac{de}{dt} \right] + |e(t)| t \right\} dt. \quad (8)$$

Here k_w is a weighting factor, $e(t)$ is the difference of the filter output signal from unity, f is a limiter function that excludes negative values, which can be specified by the expression:

$$f[x] = \max\{x, 0\}. \quad (9)$$

In [22], this coefficient is set equal to one and $k_w = 1$. These polynomials are called Chebyshev polynomials. However, the detailed studies have shown that it is very expedient to take a much larger value of this coefficient, for example, $k_w = 100$.

The coefficients of the polynomial $D_{CH}(s)$ can (and it is recommended) be found even if it is given in the simulation by the product of several polynomials [22]. For example, several polynomials of the third order, for the case when the order of the polynomial is a multiple of three. If the order of the polynomial is not a multiple of three, then some of the polynomials that are factors may have a lower order, for example, the first or the second. The paper [22] gives the results of calculating the coefficients of the polynomials $D_{CH}(s)$ up to order 12 inclusive. Although these coefficients are not given in general terms, they are only represented by numerical values. This result can be fully used in modeling or analytical calculations. In addition, in [22], the polynomials $D_{CH}(s)$ are presented as products of polynomials of lower order. This is also not a problem because, if necessary, all coefficients can be calculated by simple multiplying elementary polynomials. In addition, finding the roots of polynomials in this form is much easier, and this is irrelevant in modeling, as it is possible to use any variant of writing polynomials. It is easy to use the polynomial of the highest order among all $D_{CH}(s)$ polynomials published so far. Its analytical expression is as follows:

$$P_{CH12}(s) = (1 + a_1s + a_2s^2 + s^3)(1 + a_3s + a_4s^2 + s^3)(1 + a_5s + a_6s^2 + s^3)(1 + a_7s + a_8s^2 + s^3). \quad (10)$$

In this case, the values of the coefficients are as follows: $a_1=2.02952$, $a_2=4.36457$, $a_3=5.51585$, $a_4=5.1134$, $a_5=4.47914$, $a_6=10.5786$, $a_7=1.19969$, $a_8=3.07125$ [22]. It is noticed during modeling that the response of the filter (7) with polynomial (10) in the denominator models the response of a pure delay link with high accuracy. To demonstrate this, it is sufficient to use the *VisSim* software with the structure of the model shown in Figure 4. The simulation results are shown in Figure 5. It shows that this graph perfectly approximates the response of the delay link with a time constant equal to $\tau=14$ s.

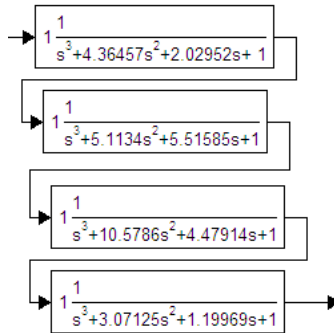


Fig. 4. Structure for modeling filter (7) with polynomial (10)

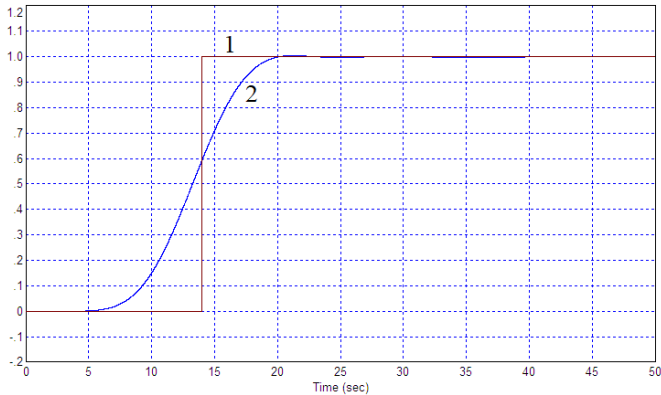


Fig. 5. Transient process in the structure according to fig. 2 (line 2) versus the ideal process (line 1)

These two graphs coincide in the initial interval from 0 to 5 s, in the final interval from 20 s and further to infinity, as well as at the point $t=14$ s. The value of the time constant $\tau = 14$ s is just found from the condition that at the moment of the jump the indicated function reaches 60% of its steady-state value, which takes place in Padé approximations of any order. Thus, filter (7), (10) can be used as a model for the delay link (1) at $\tau=14$ s. For the case of an arbitrary value of τ , it is sufficient to apply the appropriate scaling of frequencies and time. For example, if we assume that the unit of measurement on the graph in Figure 5 is the time interval equal to 14 s, then this graph corresponds to the model of the delay link with $\tau=1$ (in new units of time). Note that even the seventh-order polynomial $D_{CH}(s)$ in all parameters better approximates the delay link than approximation (6). For the case $\tau=1$, this polynomial has the following form:

$$P_{CH5}(s) = (1 + 0.4123s + 0.11493s^2 + 0.00463s^3)(1 + 0.5113s + 0.02778s^2). \quad (11)$$

Also, for comparison, the 9th order polynomial was simulated for the case $\tau=1$. This polynomial has the form:

$$P_{CH9}(s) = (1 + 0.113s + 0.13s^2 + 0.001s^3)(1 + 0.4321s + 0.02915s^2 + 0.001s^3)(1 + 0.3824s + 0.08446s^2 + 0.001s^3). \quad (12)$$

Figure 6 shows the structures for modeling the corresponding filters, and Figure 7 shows the resulting transients. It is obviously from Figure 7 that model (6) is the worst approximation of model (1) in comparison with models (11) and (12).

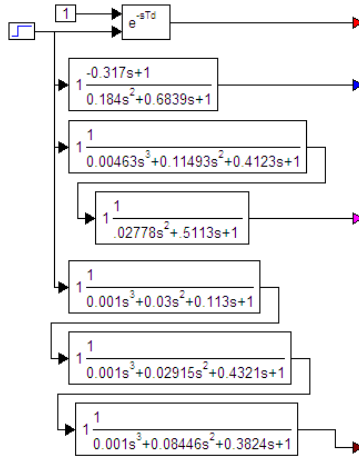


Fig. 6. Structures for modeling filters (1), (6), (11) and (12)

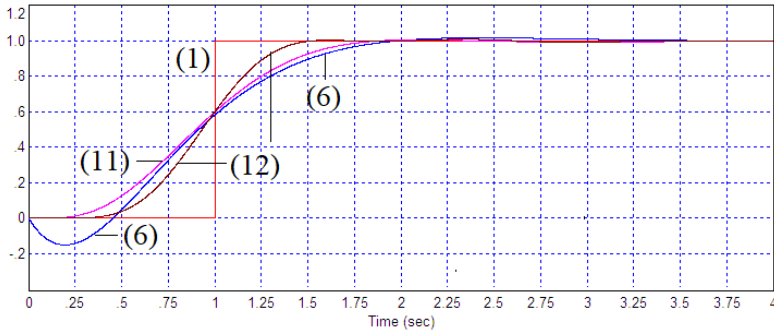


Fig. 7. Transient structures in the models according to relations (1), (6), (11) and (12), the marking of the lines corresponds to the number of the model equation

5. Control of the Object in the Form of an Integrator and a Delay Link. Using the known models (4) and (6), as well as the proposed model (10), we calculate the controller for the object with delay, after which we will simulate the system containing such a controller and the actual model of the object according to relation (1). In the same way, let us check the applicability of the Padé model and other models.

Consider an object of the form

$$W_{O1}(s) = W_M(s)W_D(s) = \frac{1}{s} e^{-\tau s}. \tag{13}$$

Let pose the problem of finding the PID controller coefficients for controlling this object in a locked loop. The regulator transfer function is:

$$W_R(s) = k_P + k_I \frac{1}{s} + k_D s. \quad (14)$$

The regulator coefficients k_P , k_I , k_D have to be calculated. We will use the numerical optimization method.

For optimization, the next experiment uses the objective (cost) function (8), (9). Initial parameter values can be set to zero. Since model (10) with the above coefficients best approximates the delay link at $\tau=14$ s, this value of τ is put in relation (10). This is acceptable since this result can always be recalculated for a different time constant by introducing a scale factor. With the new coefficient, the overall gain of the object will change. This is also insignificant since the change in the coefficient can be easily taken into account by proportional changes in all controller coefficients; therefore, problem (10) remains the most general for any arbitrary finite value of τ . The structure for designing a regulator for this object is shown in Figure 8. It fully corresponds to the structure of Figure 3 in terms of graphical programming in the VisSim software. The upper part of this figure shows the structure of the control system that will be in the implementation of this system in reality. It has a generator for the reference signal (in this case, it is a step jump), an adder through which the negative feedback is closed, a PID controller and an object. In this case, instead of an object, there is its model, which consists of a model of a delay link and an integrator. The middle and lower parts of Figure 8 show the blocks for calculating the coefficients of the controller and the calculator of the value of the cost function. As a result of the optimization, the following values of the PID controller coefficients were obtained: $k_P=0.0572943$, $k_I=-3.48581 \cdot 10^{-7}$, $k_D=0.322082$. These factors can be rounded up to 3–4 significant digits. The coefficient of the integrating path is very small, and this is natural since the object contains an integrator, therefore, the integrator is not needed in the controller. However, for the sake of the purity of the experiment, the calculated coefficients are saved without rounding. The transient process obtained in this system is shown in Figure 9.

Now the approximation (10) in the structure shown in Figure 8 is replaced by the exact model of the object (11). The obtained transient process in the new system is shown in Figure 9 (line 2), where the process obtained by the approximation is shown for comparison (line 1). It can be seen that the process in the system with a real object containing a link of a pure delay differs from the process obtained in the system with its approximation, but this difference is not significant since the overshoot does not exceed 10%.

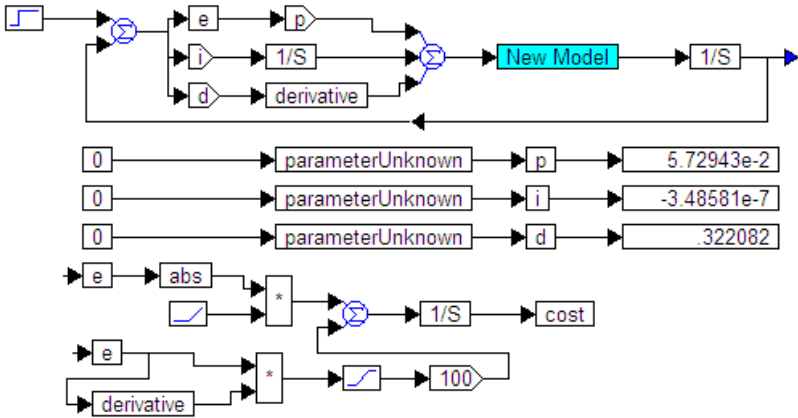


Fig. 8. Structure for designing a regulator for an object (11)

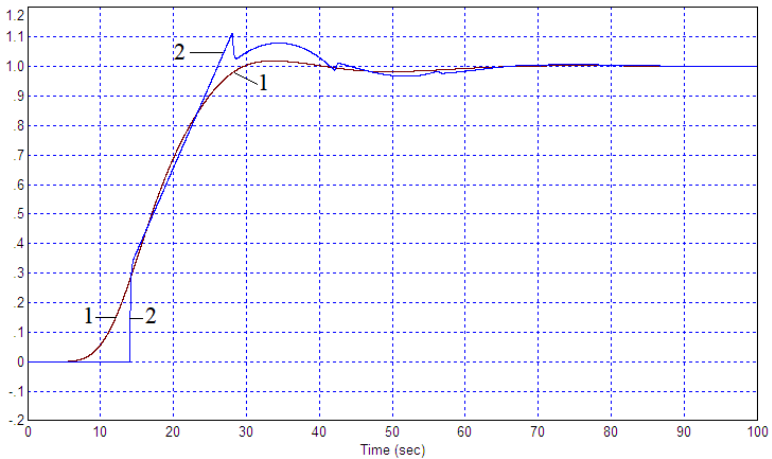


Fig. 9. Comparison of the transient process obtained using the approximation (10) (line 1) and using the exact description of the object (11) (line 2)

The comparison of the two transient processes shows that the process with the actual delay begins only after this delay time has elapsed, which is quite natural. Meanwhile, the process with approximation begins earlier, because the response in this approximate model also begins earlier, as shown in Figure 5. The response in a real system contains several small jumps, but in general, the response is quite similar to the response in a sys-

tem with an approximation. The largest difference is no more than 10%, except for the initial section, where the difference is inevitable but also not critical, as the type of the transition process in this initial section should be different in this way as well. If the Padé approximation (4) is applied in the considered and detailed method, then, in this case, the optimization with the help of this objective function will not lead to a result. The VisSim program cannot complete the calculations because the calculation results in unacceptably large values of the output signals and the objective function. The use of approximation (6) in solving this problem by the considered method also does not lead to success.

The above shows that the advantage of the proposed approximation in the form of a filter based on the polynomial $D_{CH}(s)$ model has a wider range of applicability. In particular, if the plant model consists of a delay link and an integrator, then the Padé approximation does not allow finding the PID controller by numerical methods. Some analytical methods may not give the desired result. If, in this case, a filter (7) is used as a model, then the required optimization result is achieved. In this case, the optimization procedure allows finding the optimal settings for the PID controller, with which the real system also works successfully.

Note that the use of the exact model (13) in the structure shown in Figure 8 easily leads to finding the required optimal regulator.

6. Control of the Object in the Form of an Integrator and Aperiodic Link. Consider an object of the form:

$$W_{02}(s) = \frac{1}{s+10} e^{-\tau s}. \quad (15)$$

Let's solve the same problem in the same two ways. The experiment used the fifth-order Padé approximation for $\tau=14$ s:

$$W_{1\tau}^5(s) = \frac{-s^5+30s^4-420s^3+3360s^2-15120s+30240}{s^5+30s^4+420s^3+3360s^2+15120s+30240}. \quad (16)$$

Figure 10 shows the results for calculating the optimal controller using approximation (14) (line 1) and the results of using this controller with a real object (11) (line 2). If we neglect the initial part, which must be different, then the rest of the transient processes coincide almost perfectly. The maximum difference between the actual process and the process calculated with the Padé approximation is 5%, which is a very good result because overshoot is three times less than required, it does not exceed 3%. The largest jump occurs by 5% and does not contribute to overshoot, the difference

in the process at the initial stage, when the delay is in effect, is inevitable, and it is exactly what it can be, that is, before the expiration of the delay time, there is no signal at the output, which is natural. When using the filter approximation (10), the results are approximately the same as in the case of the model with an integrator, i.e. it almost coincides with the processes shown in Figure 7. In this case, the maximum difference also reaches 10%, as for the object (10).

Consequently, the proposed model (10) has both disadvantages and advantages. The advantages have been mentioned above. The disadvantage of the model is that when using it, the obtained result coincides with less accuracy with the result obtained using the Padé model.

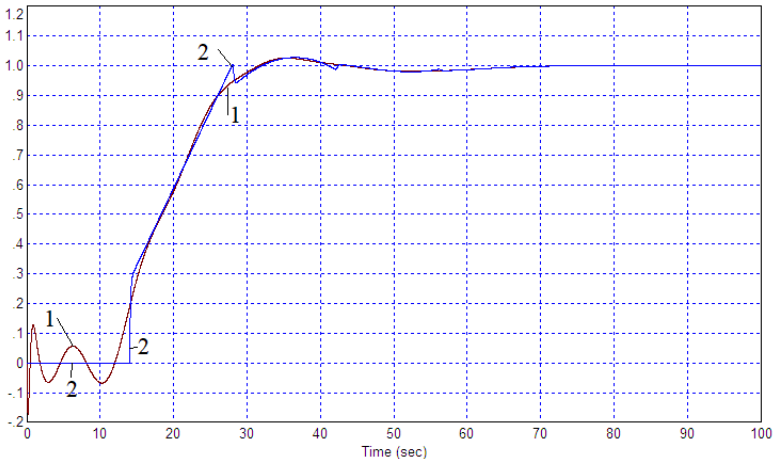


Fig. 10. Comparison of the transient process obtained using the fifth-order Padé approximation (line 1) and using the exact description of the object (11) (line 2)

7. The Use of Mixed Approximation. It was found that the serial connection of the Padé approximation model and the filter model (10) gives a much more satisfactory result. The transfer function of the composite approximation $W_C(s)$ in this case takes the following form:

$$W_C(s) = W_{CH12}(s)W_{1\tau}^5(s). \quad (17)$$

The transient process at the output of the synthetic model (17) is shown in Figure 11. It corresponds to a time delay of $\tau=28$ s.

An even better result is achieved when using the $W_{C2}(s)$ approximation in the form of a series connection of two models: the Padé approximation and filter (10) one:

$$W_{C2}(s) = W_{1\tau}^5(s)W_{CH12}(s)W_{1\tau}^5(s). \quad (18)$$

It is possible to successfully solve the object control problem of the form (13) when set $\tau=42$ s. Figure 12 shows the structure for modeling and optimization of a system with an object in form (13) using the compound approximation (18).

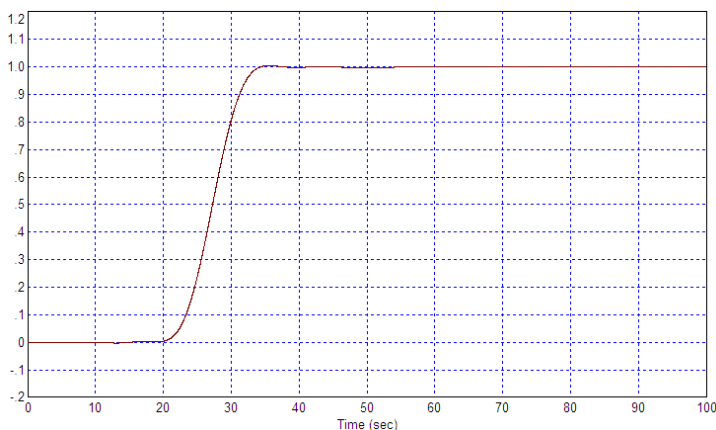


Fig. 11. Transient process at the output of the synthetic model

Figure 13 shows the transient processes calculated in the system with object (13) where approximation (16) is used instead of a pure delay (1) (line 1) and when the system has a real object (13) (line 2). Comparison of these graphs shows that the coincidence of the two processes is almost ideal with the exception of small inevitable deviations near the points $t_1=\tau$, $t_2=2\tau$, $t_3=3\tau$.

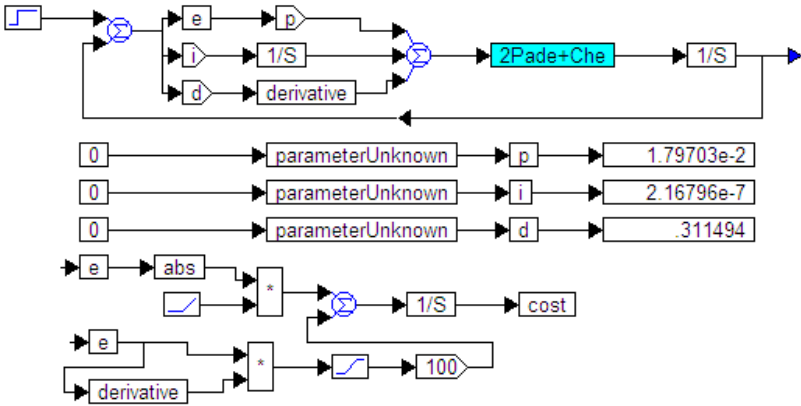


Fig. 12. Structure for modeling and optimization of a system with an object of the form (13) using compound approximation (18)

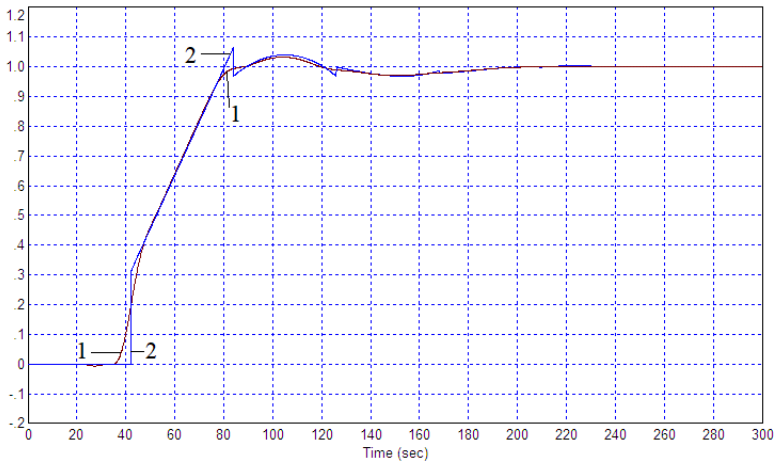


Fig. 13. Transient processes calculated in the system with approximation (18) (line 1) and in the system with a real object (10) (line 2)

Discussion. Figure 7 shows that the proposed approximation by the filter based on the polynomial (10) is better because the solution of the control problem for an object containing an integrator and a delay link can be successfully solved only by using this approximation. The rest of the known approximations, including the Padé approximation, do not work in this case when using the numerical optimization method. It follows that even when

analytical methods are used, these approximations may not be applicable or work worse than the approximation based on the polynomial (10).

Composite approximation (18) is even more accurate. The use of the two proposed approximations (10) and (18) allows for a more successful solution of the tasks for designing controllers by the methods of numerical optimization. In this case, the approximation (10) is not without drawbacks since there are models of objects, such as model (13), for which the use of the Padé approximation (4) gives more accurate results. But the advantage of approximation (10) is that there are such object models as a model according to relation (13), for which the Padé approximation does not allow solving the posed problem of regulator design, while the approximation (10) allows solving this problem effectively. Also, a composite approximation (17) and (18) is proposed, and it is much more accurate, and it allows solving all the problems considered with the highest accuracy. The accuracy of the solution of the problem was evaluated by the accuracy of the coincidence of transient processes in the resulting system using an approximation model and similar processes with the same controller, using the exact model of the object. In the case of using model (15), this accuracy is the highest one.

The proposed approximation can also be compared with other options, for example, with a filter approximation with a binomial polynomial in the denominator, or an approximation based on the Taylor series, as (3), used in publication [24], or with a polynomial of the following form [25]:

$$G(s) = \frac{1}{1+s+0.5s^2}. \quad (19)$$

All of these approximations are much less accurate and, thus, perform worse. The comparison of the integrals of the modulus of the error of these approximations with the same characteristic for the approximation based on the polynomial (10) gives the error that is 1.6 times greater.

Conclusion. The paper proposes a new approximation of the delay link with a low-pass filter. The approximation of the transcendental transfer function by a link in the form of a rational fraction of polynomials makes more efficient the use of analytical methods for designing a controller for objects with delay. The disadvantages of the Padé approximation are shown by the method of modeling and numerical optimization and three new types of approximation, free from the identified shortcomings, are proposed. Based on the research, it can be said that the known methods for approximating the delay link by a polynomial transfer function have their draw-

backs. The expediency of the most accurate approximation can take place only if it is necessary to use analytical methods of design; since numerical methods do not need such an approximation, all known packages for modeling and optimization easily simulate the exact transfer function (1). Thus, approximation by known models is far from always advisable, but only if analytical methods are used further, the known approximation methods are inferior to the proposed method.

Calculations of the proposed approximation up to the ninth order have been carried out. Increasing the order leads to a more accurate approximation of the stepwise delay jump. Figure 14 shows transient processes in filters of the fifth and sixth order, Figure 15 shows the processes in filters of the eighth and ninth order.

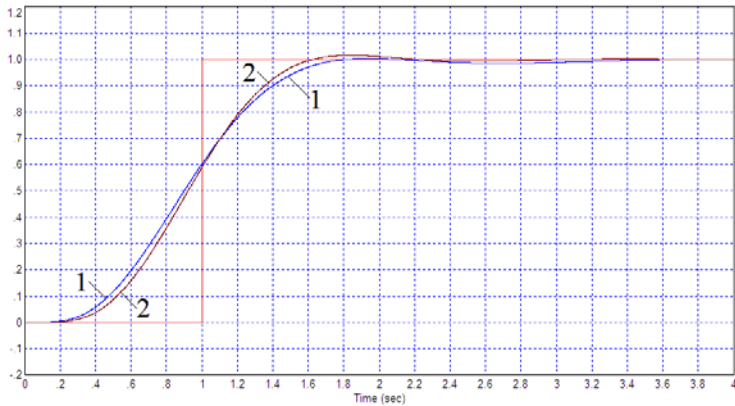


Fig. 14. Transient processes in filters of the fifth order (line 1) and sixth order (line 2) according to the proposed method

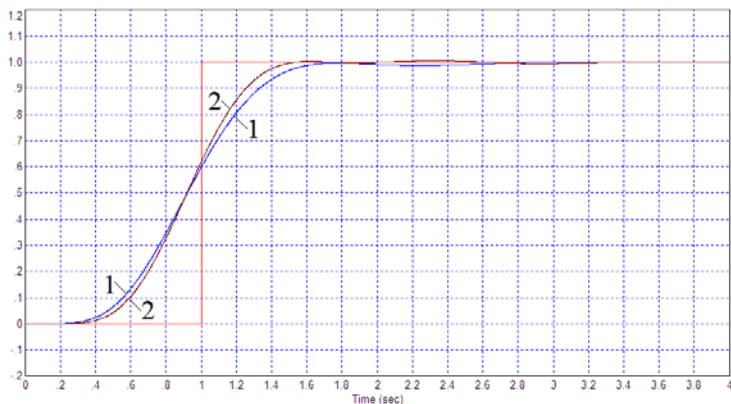


Fig. 15. Transient processes in filters of the 8th order (line 1) and 9th order (line 2) according to the proposed method

Raising the order above the fifth does not give a noticeable advantage in optimization, since fairly reliable results are obtained even when using models of the 3rd – 5th order. Therefore, such an excessive complication of the model by increasing the order of approximation is apparently inappropriate.

References

1. Bušek J., Zítek P., Vyhřídál T. Astatism analysis of time delay controllers towards effective anti-windup schemes. 2019 22nd International Conference on Process Control (PC19), 2019, pp. 74-79, doi: 10.1109/PC.2019.8815283.
2. Cao M. and Yang J. The Effect of the Approximation Method for Large Time Delay Process on the Performance of IMC-PID Controller. 2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCT), 2018, pp. 73-77, doi: 10.1109/ICCPCT.2018.8574299.
3. Tyutikov V. V., Voronenkova A. A. Analytical synthesis and analysis of industrial facility control system versions. 2017 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2017, pp. 1-5, doi: 10.1109/ICIEAM.2017.8076118.
4. Zhang K., Li K. Quantitative robust control design of discharge air temperature system. 2017 11th Asian Control Conference (ASCC), 2017, pp. 2612-2617, doi: 10.1109/ASCC.2017.8287588.
5. Muresan C. I., Birs I. R., Darab C., Prodan O., Keyser R. De. Alternative Approximation Method for Time Delays in an IMC Scheme. 2019 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2019 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 2019, pp. 532-539, doi: 10.1109/ACEMP-OPTIM44294.2019.9007220.
6. Tyutikov V. V., Voronenkova A. A. Analytical synthesis and analysis of industrial facility control system versions. 2017 International Conference on Industrial

- Engineering, Applications and Manufacturing (ICIEAM), 2017, pp. 1-5, doi: 10.1109/ICIEAM.2017.8076118.
7. Xing H., Ploeg J., Nijmeijer H. Padé Approximation of Delays in Cooperative ACC Based on String Stability Requirements. Published. 2016. Computer Science, Mathematics. IEEE Transactions on Intelligent Vehicles. URL: <https://www.semanticscholar.org/paper/Pad%C3%A9-Approximation-of-Delays-in-Cooperative-ACC-on-Xing-Ploeg/949b761f628c2d4156b14211c5ec62b714070d56?sort=relevance&citationIntent=methodology> (accessed 03.09.2021).
 8. Sitnikov E.A. Mathematical modeling and development of delay compensation models for control systems of polymerization processes. Dissertation for the degree of candidate of technical sciences in specialty 05.13.18. 2004 URL: <https://www.dissercat.com/content/matematicheskoe-modelirovanie-i-razrabotka-modelei-kompensatsii-zapazdyvaniya-dlya-sistem-up> (accessed 03.09.2021).
 9. Tkhan V.Z. Synthesis of automatic control systems with delay by the numerical method. Dissertation for the degree of candidate of technical sciences in specialty 05.13.01. 2018 URL: https://etu.ru/assets/files/nauka/dissertacii/2018/zung/avtoreferat_zung.pdf (accessed 03.09.2021).
 10. Karimov V.S. Synthesis of automatic control systems for multiply connected objects with delays on the basis of system embedding technology. Dissertation for the degree of candidate of technical sciences in the specialty. 05.13.01. URL: https://ugatu.su/media/uploads/MainSite/Science/dissovet/03/2012/11.03.13/karimov_avtoreferat.pdf (accessed 03.09.2021).
 11. Kushwaha B. K., Narain A. Controller design for Cuk converter using model order reduction. 2012 2nd International Conference on Power, Control and Embedded Systems, 2012, pp. 1–5.
 12. Jedynak R. B., Gilewicz J. Computation of the c-Table Related to the Padé Approximation. Hindawi Publishing Corporation Journal of Applied Mathematics V. 2013, Article ID 185648, 10 p. URL: <http://dx.doi.org/10.1155/2013/185648>
 13. Turut O. V., Güzel N. Multivariate Padé Approximation for Solving Nonlinear Partial Differential Equations of Fractional. Hindawi Publishing Corporation Abstract and Applied Analysis V. 2013, Article ID 746401, 12 p.
 14. Bhattacharjee S., Banerjee A., Neogi B. An Application of Pade Approximation and PID Tuning Technique to Improve the System Performance of Electric Ventricular Assist Device. International Journal of Advanced Scientific Research and Management, V. 4 Issue 5, May 2019. pp. 274–280. URL: http://ijasrm.com/wp-content/uploads/2019/05/IJASRM_V4S5_1467_274_280.pdf (accessed 03.09.2021).
 15. Ognev S.P. Variant approximations of the delay link. <https://www.sworld.education/konfer24/142.htm> (accessed 03.09.2021).
 16. Stopakevich A.A., Stopakevich A.A. Stable precision modeling of the delay operator. URL: <https://biblio.onat.edu.ua/bitstream/handle/123456789/3081/Stopakevych.pdf?sequence=1&isAllowed=y> (accessed 03.09.2021). (In Russ.)
 17. Truntyagin I.M. Analysis of the stability of automatic control systems with delay using the Bode integral // Synergy of Sciences. 2017. No. 18. pp. 888–899. URL: <http://synergy-journal.ru/archive/article1516> (accessed 03.09.2021). (In Russian)
 18. The Padé Approximation and its Physical Applications. January. 1972. URL: <https://www.researchgate.net/publication/229531069> The Pade Approximation and its Physical Applications (accessed 03.09.2021).

19. Morel J.-M., Teissier B. Pade Approximation and its Applications. Proceedings of a Conference held in Antwerp, Belgium, 1979. Lecture Notes in Mathematics. URL: <https://www.springer.com/series/304>. (accessed 03.09.2021).
20. Baker J., Graves-Morris P. Padé Approximations. Cambridge Univ. Press, 1996, 764 p.
21. Tkhan V.Z., Dementiev Yu.N., Goncharov V.I. Improving the accuracy calculation of time delay automatic control. Software & Systems (Programmnye Produkty i Sistemy). 2018, V. 31, N3. pp. 521 – 526. URL: <http://swsys.ru/index.php?page=article&id=4495> (accessed 03.09.2021). (In Russ.)
22. Zhmud V.A., Dimitrov L., Nosek J. Automatic Control Systems. New Concepts and Structures of Regulators. Monograph. Moscow: RuScience, 2018. 84 p.
23. Zhmud V.A., Dimitrov L.V. Choice of High-Order Characteristic Polynomials of Automatic Control of Closed System. Automatics & Software Engineering. 2016. 2(16). pp.35–46
URL:http://jurnal.nips.ru/sites/default/files/%D0%90%D0%98%D0%9F%D0%98-2-2016-5_1.pdf (accessed 03.09.2021). (In Russ.)
24. Qibing, J., Ling, Q., Xuewei, W., Fei, Q. (2009, August). Base on all-pole approximation a new internal model PID control method for the system with time delays. In 2009 International Conference on Mechatronics and Automation (pp. 268-273). IEEE. URL: <https://doi.org/10.1109/ICMA.2009.5246112> (accessed 01.12.2021).
25. Dudarenko, N.A., Polinova, N.A., Serzhantova, M.V., & Ushakov, A.V. (2014). Multiple binomial structures in the problem of approximating dynamic chains containing a pure delay link. Proceedings of higher educational institutions. Instrumentation (Priborostroenie). 57(7). URL: https://openbooks.itmo.ru/read_pribor/10258/10258.pdf (accessed 01.12.2021). (In Russ.)

Zhmud Vadim — Ph.D., Dr.Sci., Associate Professor, Leading researcher, Institute of Laser Physics SB RAS. Research interests: automation, electronics, optoelectronics, robotics, measuring technology, automatic control theory, numerical modeling and optimization, radio physics, laser physics, information technology and computing, software and systems. The number of publications — 600. oao_nips@bk.ru; 10, Karl Marx Avenue, 630090, Novosibirsk, Russia; office phone: +7(913)473-2997.

Dimitrov Lubomir — Ph.D., Dr.Sci., Professor, Vice-rector, Mechanical engineering faculty, department of machine elements and nonmetallic constructions, Technical University of Sofia. Research interests: design in mechanical engineering and mechatronics, robotics and automatic control systems. The number of publications — 204. lubomir_dimitrov@tu-sofia.bg; 8, Boulevard "St. Kliment Ohridski", 1756, Sofia, Bulgaria; office phone: +3(592)965-2111.

Sablina Galina — Ph.D., Associate professor, Novosibirsk State Technical University. Research interests: automation, electronics, automatic control theory, numerical modeling and optimization. The number of publications — 150. sablina@corp.nstu.ru; 10, Karl Marx Avenue, 630073, Novosibirsk, Russia; office phone: +7(383)346-1119.

Roth Hubert — Ph.D., Dr.Sci., Professor, Head of the department of automatic control engineering, University of Siegen. Research interests: automatic control engineering, computers, telematics, automatic, control. The number of publications — 200. hubert.roth@uni-siegen.de; 2, Adolf-Reichwein-Strasse, 57076, Siegen, Germany; office phone: +492717400.

Nosek Jaroslav — Ph.D., Dr.Sci., Professor, Institute of mechatronics and computer engineering, Technical University of Liberec. Research interests: area of ferroelectric thin films and their integration into microelectromechanical systems, control of mechatronic systems. The number of publications — 90. jaroslav.nosek@tul.cz; 1402/2, Student St., 46117, Liberec, Czechia; office phone: +420-485-351-111.

Hardt Wolfram — Ph.D., Dr.Sci., Vice-dean on international affairs, Technical University of Chemnitz. Research interests: technical informatics. The number of publications — 350. hardt@cs.tu-chemnitz.de; 62, St. of Nations, 09111, Chemnitz, Germany; office phone: +49 371 531-0.

В.А. ЖМУДЬ, Л. ДИМИТРОВ, Г.В. САБЛИНА, Г. РОТ, Я. НОСЕК, В. ХАРДТ
**О ЦЕЛЕСООБРАЗНОСТИ И ВОЗМОЖНОСТЯХ
АППРОКСИМАЦИИ ЗВЕНА С ЧИСТЫМ ЗАПАЗДЫВАНИЕМ**

Жмудь В.А., Димитров Л., Саблина Г.В., Рот Г., Носек Я., Хардт В. О целесообразности и возможностях аппроксимации звена с чистым запаздыванием.

Аннотация. При решении задач управления объектом с запаздыванием часто необходимо аппроксимировать звено чистого запаздывания минимально фазовым звеном, чтобы обеспечить возможность использования аналитических методов для проектирования регулятора. Существует множество методов аппроксимации, основанных на разложении в ряд Тейлора, а также модифицированных методов. Наиболее известен метод аппроксимации Паде. Известные методы аппроксимации имеют существенные недостатки, которые выявляет данная работа. Однако существуют и другие способы формирования других типов фильтров, которые могут служить лучшим приближением при определении соотношения задержек, хотя они и не используются для этих целей. В частности, известны способы формирования искомого дифференциального уравнения замкнутой системы заданного порядка методом численной оптимизации. В этом случае замкнутая система ведет себя как фильтр соответствующего порядка, числитель которого равен единице, а указанный полином стоит в знаменателе. Моделирование показало, что такой фильтр является эффективной альтернативной аппроксимацией звена задержки и может использоваться для тех же целей, для которых предполагалось использовать аппроксимацию Паде. Полиномиальные коэффициенты в литературе рассчитывались только до 12-го порядка. Чем выше порядок полинома, тем точнее аппроксимация.

Ключевые слова: формула Паде, запаздывание, аппроксимация, управление, автоматизация.

Жмудь Вадим Аркадьевич — д-р техн. наук, профессор, главный научный сотрудник, Институт лазерной физики СО РАН. Область научных интересов: автоматика, электроника, оптоэлектроника, робототехника, измерительная техника, теория автоматического управления, численное моделирование и оптимизация, радиофизика, лазерная физика, информационные технологии и вычислительная техника, программное обеспечение и системы. Число научных публикаций — 600. oao_nips@bk.ru; проспект Карла Маркса, 10, 630090, Новосибирск, Россия; р.т.: +7(913)473-2997.

Димитров Любомир — д-р техн. наук, профессор, проректор, машиностроительный факультет, кафедра элементов машин и неметаллических конструкций, Софийский технический университет. Область научных интересов: проектирование в машиностроении и мехатронике, робототехнике и системах автоматического управления. Число научных публикаций — 204. lubomir_dimitrov@tu-sofia.bg; бульвар "св. Климент Охридски", 8, 1756, София, Болгария; р.т.: +3(592)965-2111.

Саблина Галина Владимировна — канд. техн. наук, доцент, Новосибирский государственный технический университет (НГТУ). Область научных интересов: автоматика, электроника, теория автоматического управления, численное моделирование и оптимизация. Число научных публикаций — 150. sablina@corp.nstu.ru; проспект Карла Маркса, 10, 630073, Новосибирск, Россия; р.т.: +7(383)346-1119.

Рот Губерт — д-р техн. наук, профессор, заведующий кафедрой техники автоматического управления, Университет Зигена. Область научных интересов: техника автоматического управления, компьютеры, телематика, автоматика, управление. Число научных публикаций — 200. hubert.roth@uni-siegen.de; Адольф-Райхвайн-Штрассе, 2, 57076, Зиген, Германия; р.т.: +492717400.

Носек Ярослав — д-р техн. наук, профессор, институт мехатроники и вычислительной техники, Технический университет Либереца. Область научных интересов: область тонких сегнетоэлектрических пленок и их интеграция в микроэлектромеханические системы, управление мехатронными системами. Число научных публикаций — 90. jarooslav.nosek@tul.cz; Студенческая, 1402/2, 46117, Либерец, Czechia; р.т.: +420-485-351-111.

Хардт Вольфрам — д-р техн. наук, заместитель декана по международным связям, Технический университет Хемница. Область научных интересов: техническая информатика. Число научных публикаций — 350. hardt@cs.tu-chemnitz.de; Наций, 62, 09111, Хемниц, Германия; р.т.: +49 371 531-0.

Литература

1. Bušek J., Zitek P., Vyhřídál T. Astatism analysis of time delay controllers towards effective anti-windup schemes. 2019 22nd International Conference on Process Control (PC19), 2019, pp. 74-79, doi: 10.1109/PC.2019.8815283.
2. Cao M. and Yang J. The Effect of the Approximation Method for Large Time Delay Process on the Performance of IMC-PID Controller. 2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCT), 2018, pp. 73-77, doi: 10.1109/ICCPCT.2018.8574299.
3. Tyutikov V. V., Voronenkova A. A. Analytical synthesis and analysis of industrial facility control system versions. 2017 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2017, pp. 1-5, doi: 10.1109/ICIEAM.2017.8076118.
4. Zhang K., Li K. Quantitative robust control design of discharge air temperature system. 2017 11th Asian Control Conference (ASCC), 2017, pp. 2612-2617, doi: 10.1109/ASCC.2017.8287588.
5. Muresan C. I., Birs I. R., Darab C., Prodan O., Keyser R. De. Alternative Approximation Method for Time Delays in an IMC Scheme. 2019 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2019 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 2019, pp. 532-539, doi: 10.1109/ACEMP-OPTIM44294.2019.9007220.
6. Tyutikov V.V., Voronenkova A. A. Analytical synthesis and analysis of industrial facility control system versions. 2017 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), 2017, pp. 1-5, doi: 10.1109/ICIEAM.2017.8076118.
7. Xing H., Ploeg J., Nijmeijer H. Padé Approximation of Delays in Cooperative ACC Based on String Stability Requirements. Published. 2016. Computer Science, Mathematics. IEEE Transactions on Intelligent Vehicles. URL: <https://www.semanticscholar.org/paper/Pad%C3%A9-Approximation-of-Delays-in-Cooperative-ACC-on-Xing-Ploeg/949b761f628c2d4156b14211c5ec62b714070d56?sort=relevance&citationIntent=methodology> (accessed 03.09.2021).
8. Sitnikov E.A. Mathematical modeling and development of delay compensation models for control systems of polymerization processes. Dissertation for the degree of

- candidate of technical sciences in specialty 05.13.18. 2004 URL: <https://www.dissercat.com/content/matematiceskoe-modelirovanie-i-razrabotka-modelei-kompensatsii-zapazdyvaniya-dlya-sistem-up> (accessed 03.09.2021).
9. Tkhan V.Z. Synthesis of automatic control systems with delay by the numerical method. Dissertation for the degree of candidate of technical sciences in specialty 05.13.01. 2018 URL: https://etu.ru/assets/files/nauka/dissertacii/2018/zung/avtoreferat_zung.pdf (accessed 03.09.2021).
 10. Karimov V.S. Synthesis of automatic control systems for multiply connected objects with delays on the basis of system embedding technology. Dissertation for the degree of candidate of technical sciences in the specialty. 05.13.01. URL: https://ugatu.su/media/uploads/MainSite/Science/dissovet/03/2012/11.03.13/karimov_avtoreferat.pdf (accessed 03.09.2021).
 11. Kushwaha B. K., Narain A. Controller design for Cuk converter using model order reduction. 2012 2nd International Conference on Power, Control and Embedded Systems, 2012, pp. 1–5.
 12. Jedynek R. B., Gilewicz J. Computation of the c-Table Related to the Padé Approximation. Hindawi Publishing Corporation Journal of Applied Mathematics V. 2013, Article ID 185648, 10 p. URL: <http://dx.doi.org/10.1155/2013/185648>
 13. Turut O. V., Güzel N. Multivariate Padé Approximation for Solving Nonlinear Partial Differential Equations of Fractional. Hindawi Publishing Corporation Abstract and Applied Analysis V. 2013, Article ID 746401, 12 p.
 14. Bhattacharjee S., Banerjee A., Neogi B. An Application of Pade Approximation and PID Tuning Technique to Improve the System Performance of Electric Ventricular Assist Device. International Journal of Advanced Scientific Research and Management, V. 4 Issue 5, May 2019. pp. 274–280. URL: http://ijasrm.com/wp-content/uploads/2019/05/IJASRM_V4S5_1467_274_280.pdf (accessed 03.09.2021).
 15. Ognev S.P. Variant approximations of the delay link. <https://www.world.education/konfer24/142.htm> (accessed 03.09.2021).
 16. Stopakevich A.A., Stopakevich A.A. Stable precision modeling of the delay operator. URL: <https://biblio.onat.edu.ua/bitstream/handle/123456789/3081/Stopakevych.pdf?sequence=1&isAllowed=y> (accessed 03.09.2021). (In Russ.)
 17. Truntyagin I.M. Analysis of the stability of automatic control systems with delay using the Bode integral // Synergy of Sciences. 2017. No. 18. pp. 888–899. URL: <http://synergy-journal.ru/archive/article1516> (accessed 03.09.2021). (In Russian)
 18. The Padé Approximation and its Physical Applications. January. 1972. URL: <https://www.researchgate.net/publication/229531069> The Pade Approximation and its Physical Applications (accessed 03.09.2021).
 19. Morel J.-M., Teissier B. Pade Approximation and its Applications. Proceedings of a Conference held in Antwerp, Belgium, 1979. Lecture Notes in Mathematics. URL: <https://www.springer.com/series/304>. (accessed 03.09.2021).
 20. Baker J., Graves-Morris P. Padé Approximations. Cambridge Univ. Press, 1996, 764 p.
 21. Tkhan V.Z., Dementiev Yu.N., Goncharov V.I. Improving the accuracy calculation of time delay automatic control. Software & Systems (Programmye Produkty i Sistemy). 2018, V. 31, N3. pp. 521 – 526. URL: <http://swsys.ru/index.php?page=article&id=4495> (accessed 03.09.2021). (In Russ.)
 22. Zhmud V.A., Dimitrov L., Nosek J. Automatic Control Systems. New Concepts and Structures of Regulators. Monograph. Moscow: RuScience, 2018. 84 p.
 23. Zhmud V.A., Dimitrov L.V. Choice of High-Order Characteristic Polynomials of Automatic Control of Closed System. Automatics & Software Engineering. 2016. 2(16).

pp.35–46)

URL:http://jurnal.nips.ru/sites/default/files/%D0%90%D0%98%D0%9F%D0%98-2-2016-5_1.pdf (accessed 03.09.2021). (In Russ.)

24. Qibing, J., Ling, Q., Xuwei, W., Fei, Q. (2009, August). Base on all-pole approximation a new internal model PID control method for the system with time delays. In 2009 International Conference on Mechatronics and Automation (pp. 268-273). IEEE. URL: <https://doi.org/10.1109/ICMA.2009.5246112> (accessed 01.12.2021).
25. Dudarenko, N.A., Polinova, N.A., Serzhantova, M.V., & Ushakov, A.V. (2014). Multiple binomial structures in the problem of approximating dynamic chains containing a pure delay link. Proceedings of higher educational institutions. Instrumentation (Priborostroenie). 57(7). URL: https://openbooks.itmo.ru/read_pribor/10258/10258.pdf (accessed 01.12.2021). (In Russ.)

С.С. МАНАЕНКО, С.В. ДВОРНИКОВ, А.В. ПШЕНИЧНИКОВ
**ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ФОРМИРОВАНИЯ
СИГНАЛЬНЫХ КОНСТРУКЦИЙ СЛОЖНОЙ СТРУКТУРЫ**

Манаенко С.С., Дворников С.В., Пшеничников А.В. Теоретические аспекты формирования сигнальных конструкций сложной структуры.

Аннотация. Приемно-передающие тракты современных систем радиосвязи строятся на основе открытой структуры, обеспечивающей иерархическое разграничение доступа к предоставляемым телекоммуникационным услугам. Однако такой подход не исключает возможность доступа к передаваемому контенту со стороны несанкционированных пользователей. Скрытие информации методами криптографической защиты в такой ситуации только активизирует дополнительный интерес к передаче, поэтому наиболее прагматичным решением является использование сигналов сложной структуры, существенно затрудняющих или даже исключающих извлечение из них информации сторонними пользователями. Проблема регламентирования селекции доступа при разработке и проектировании элементов радиосистем является достаточно многогранной и обладает высокой степенью сложности. Одно из направлений решения задач данной предметной области базируется на известных подходах расширения базы сигналов, однако алгоритмы их практической реализации получены без учета ограничений на выделенный ресурс и без учета самого факта применения этих алгоритмов. На основе теории систем и общей теории связи разработан подход к формированию сигнальных конструкций сложной структуры, обеспечивающий повышение свойств их структурной скрытности по отношению к несанкционированным пользователям. При этом были уточнены известные решения на физическом уровне сигнальных пространств, что позволило формализовать процедуры формирования радиосигналов, обладающими заданными свойствами. Обоснован метод формализации функции отображения пространства сигналов на основе выделения стохастических свойств псевдослучайных последовательностей, что позволило обеспечить неопределенность их структуры при несанкционированной обработке. Апробация предлагаемого подхода приведена на примере формирования сигналов квадратурной модуляции с учетом последующего анализа их свойств с различных позиций легитимных и нелегитимных пользователей. Полученные результаты подтверждают возникшую неопределенность при нелегитимной обработке при незначительном ухудшении свойств помехоустойчивости систем радиосвязи. В целом это позволяет сделать вывод об адекватности теоретических решений. В качестве примера представлены констеляционные диаграммы сигналов на выходе квадратурного приемника. Представленная в работе совокупность предлагаемых технических решений определяет новизну данного подхода. Решаемая научная задача относится к классу задач синтеза сигналов сложных структур.

Ключевые слова: сигналы сложной структуры, структурная скрытность сигналов, энергетическая скрытность сигналов, синтез сигналов сложной структуры, сигнальные конструкции.

1. Введение. Электромагнитная доступность к сигналам систем радиосвязи (СРС) в процессе передачи сообщений со стороны несанкционированных пользователей обуславливает необходимость разработки и последующей реализации различных ограничивающих меро-

приятый как организационного, так и технического характера.

Данная проблематика многогранна, и достаточно часто, например в [1], рассматривается с позиций регламентирования селективного доступа к передаваемой (транслируемой) информации.

В других источниках, в частности в [2], электромагнитная доступность радиоканала вообще рассматривается с позиций составной части средств информационной войны.

Однако сложность решения задачи, связанной со снижением электромагнитной доступности к радиоканалам для несанкционированных абонентов, вынуждает пользователей СРС использовать криптографические меры скрытия информации. В частности, такой подход к защите информационного обмена предложен в [3], где предложено использовать различные комбинации симметричных и ассиметричных криптоалгоритмов. Данное направление достаточно активно развивается, о чем свидетельствуют многочисленные работы в данном направлении [4-6].

Вместе с тем сам факт использования методов криптозащиты выступает в качестве явного признака желания сокрытия информации, что может вызвать дополнительный интерес со стороны несанкционированных абонентов к содержанию передаваемого информационного контента [2]. Поэтому прагматичным в этом плане видится использование методов снижения информационной доступности передаваемой информации для несанкционированных абонентов за счет повышения структурной скрытности передаваемых сигналов, предложенных в [7].

В настоящее время данное направление активно развивается, о чем свидетельствуют многочисленные работы по этой проблематике.

Так, в [8] предлагается в СРС применять стохастические параллельно-последовательные широкополосные сигналы, обладающие свойствами высокой энергетической и структурной скрытности. В [9-10] для СРС разработаны модели передачи сообщений на основе хаотических сигналов, обладающих высокой структурной скрытностью.

Модели комбинированных сигнальных конструкций с регулируемым уровнем их структурной и энергетической скрытности разработаны в [11-12].

Следует отметить, что рассмотренное направление повышения скрытности связано с существенным расширением спектра сигналов, что достаточно сложно обеспечить в условиях ограничений на полосу частот, выделяемую для радиоканалов в СРС. Вместе с тем достаточно интересный подход обоснован в [13-15]. Авторы предлагают формировать сигналы таким образом, чтобы в трактах приема несанкционированных абонентов их параметры соответствовали структурам совсем

других конструкций. Очевидно, что такой подход предполагает сложность синтеза сигнальных конструкций СРС, обусловленную высокой вариативностью реальной сигнально-помеховой обстановки. Поэтому для ее учета процессы, описывающие условия функционирования СРС, наделяют широким диапазоном стохастических свойств [16-17]. При этом наибольшую многозначность варианты решения задач данного класса приобретают в условиях деструктивных воздействий, стратегии которых в явном виде определяются с недостаточной степенью точности [18]. Для таких ситуаций не определены единые подходы к постановке и решению проблемы разработки сигналов СРС, что связано со сложностью их формализации на различных уровнях эталонной модели взаимодействия открытых систем.

В соответствии с указанными обстоятельствами в настоящей работе предложен вариант разработки теоретического подхода к представлению и формированию сигналов СРС, обладающих свойствами высокой структурной и энергетической скрытности по отношению к несанкционированным абонентам (далее они определены как сигналы сложных структур) с позиций сложных систем в терминах общей теории связи и статистической радиотехники.

2. Анализ подходов к формированию сигналов сложной структуры с позиций общей теории систем. Согласно [19] к сигналам сложной структуры относят сигналы, помехоустойчивые свойства которых определяются их базой. Как правило, их синтез осуществляется различными методами расширения спектра.

Наиболее известные решения базируются на способе прямого расширения спектра путем непосредственного перемножения псевдослучайных последовательностей (ПСП) с выборками высокочастотных колебаний. В результате указанных операций получают фазоманипулированные широкополосные сигналы [19-22].

Другой подход к синтезу сигналов сложной структуры связан со стохастическим управлением частотным ресурсом СРС, который позволяет формировать сигналы с псевдослучайным законом перестройки рабочих частот (ППРЧ) [23-26].

Заметим, что методы формирования сигналов на основе ПСП предъявляет достаточно высокие требования к модели канала радиосвязи и параметрам чувствительности радиоприемных устройств [23].

Указанное обстоятельство существенно ограничивает их практическое применение, а в условиях деструктивных воздействий дополнительно усложняет модели СРС [27], поскольку в этом случае необходимо решать задачи ковариационных и корреляционных вычислений [28].

Кроме того, сигналы этого класса имеют достаточно сложную структуру и изначально предполагают наличие широкополосных каналов с однородной структурой [19].

Сигналы с ППРЧ априори предъявляют высокие требования к системе управления СРС [23]. Однако применение методов широкополосного частотного синтеза фактически нивелирует их свойства, характерные для сигналов со сложной структурой [29].

Кроме того, следует отметить, что известные в общей теории связи аналитические решения по формированию помехоустойчивых сигналов [30-31] не учитывают всю полноту возможных стратегий деструктивных воздействий, поскольку изначально предполагают в качестве помех только аддитивный белый гауссов шум (АБГШ).

Вместе с тем в [32] аргументировано, что адекватный учет деструктивных воздействий когнитивного характера возможен только при детализации радиоэлектронного конфликта методами теории управления и общей теории систем. Но такой подход предполагает представление сигналов СРС как результат формализации моделей сложными структурами [33].

С учетом указанных обстоятельств рассмотрим структуру СРС [34], как сложную систему, определяющую стратегию формирования сигналов с позиций увеличения их параметрического пространства (рис. 1).

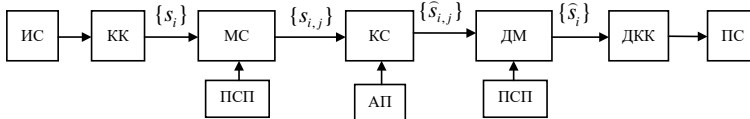


Рис. 1. Структурная схема системы связи

В общем случае, структуру СРС определяют следующие функциональные процессы.

Формирование из стохастической информации источника сообщения (ИС) электрических сигналов, которые в зависимости от способа обработки могут быть непрерывными или дискретными.

В том случае, когда ИС формирует непрерывные сигналы, то первой функцией кодера источника является аналого-цифровое преобразование, и лишь затем реализуются функции согласования алфавита источника сообщения с алфавитом канального кодера (КК) и устранение избыточности передаваемых сообщений [34].

Канальное кодирование повышает достоверность передаваемых

сообщений за счет введения избыточности. На выходе КК для построения сигнальных конструкций цифровой поток разбивается на блоки, формируя конечное множество сигналов $S^M = \{s_i\}$, $i = 0, 1, \dots, M - 1$.

Модулятор сигналов (МС) под управлением генератора ПСП расширяет пространство существования сигналов M до уровня $N = M \times L$, где L – размерность чисел, формируемых генератором ПСП. Затем модулятор генерирует гармоническое несущее колебание $S^N = \{s_{i,j}\}$, $i = 0, 1, \dots, M - 1$, $j = 0, 1, \dots, L - 1$ параметры которого изменяются в зависимости от структуры сигнала из сформированного пространства N .

Причем указанные пространства сигналов должны удовлетворять следующим свойствам [35-36]:

1) пространство сигналов $S^M = \bigcup_{i=0}^{M-1} s_i = \{s_i\}$ является метрическим и линейным;

2) d_{\min}^M – минимальное евклидово расстояние в пространстве сигналов $S^M = \{s_i\}$;

3) i -й сигнал пространства $S^M = \{s_i\}$ отображается на пространство $S_i^N = \bigcup_{j=0}^{L-1} s_{i,j} = \{s_{i,j}\}$, т.е. $S_i^N = f_{L-1}(s_i)$; область определения отображения f_{L-1} является сигнал s_i , а область изображений является пространство $S_i^N = \{s_{i,j}\}$, включающее в себя L образов сигнала s_i ;

4) в результате отображения f_{L-1} пространство сигналов $S^M = \{s_i\}$ разбивается на $N = M \times L$ элементов, т.е. $S^N = \bigcup_{i=0}^{M-1} S_i^N = \bigcup_{i=0}^{M-1} \bigcup_{j=0}^{L-1} s_{i,j} = \{s_{i,j}\}$, при этом $S_i^N \cap S_j^N = \emptyset$ для $i \neq j$;

5) d_{\min}^N – минимальное евклидово расстояние в пространстве сигналов $S^N = \{s_{i,j}\}$, при этом $d_{\min}^N < d_{\min}^M$;

6) пространства сигналов $S^M = \{s_i\}$ и $S_i^N = \{s_{i,j}\}$ являются взаимно-однозначными.

Сущность процессов преобразования сигналов в СРС поясняет на рисунке 2. Системной операцией при преобразованиях сигналов

является функция отображения f_{L-1} , которая определяет порядок перехода пространства сигналов $S^M = \{s_i\}$ в пространство $S^N = \{s_{i,j}\}$. Данная функция должна обладать следующими свойствами.

Соответствия, согласно которому каждый из сигналов s_i и $s_{i,j}$ имеет свое положение в пространстве, которое характеризуется соответствующими координатами, поскольку пространства сигналов S^M и S^N являются метрическими и линейными.

Размерности, в соответствии с которым пространства сигналов $S^M = \{s_i; d_{\min}^M\}$ и $S^N = \{s_{i,j}; d_{\min}^N\}$ (где $(M < N)$) отличаются минимальным евклидовым расстоянием ($d_{\min}^M > d_{\min}^N$).

Упорядочения, определяющего то, что отображение пространства сигналов S^M на пространство S^N будет заключаться в разбиении пространства сигналов S^M на непересекающиеся пространства S_i^N , при котором увеличивается количество сигналов до N и, соответственно, уменьшается расстояние между сигналами.

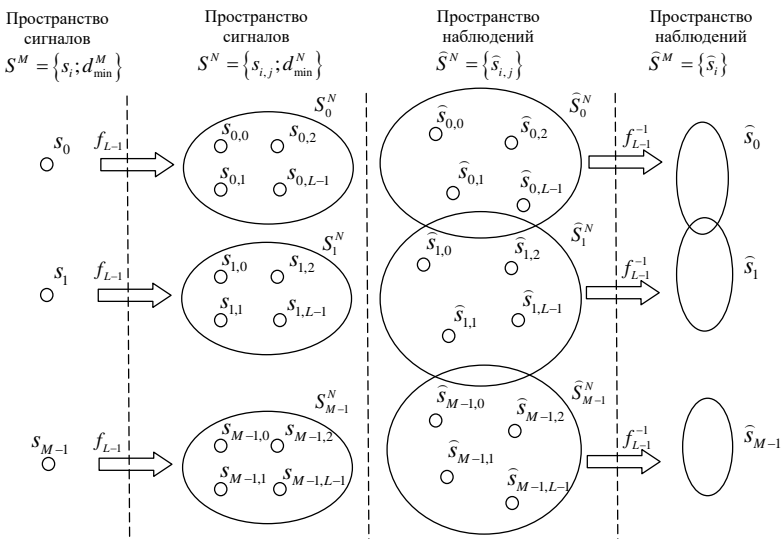


Рис. 2. Процесс преобразований сигналов в CPC

Особенностью функции отображения f_{L-1} является то, что она формализуется методами выделения стохастических свойств на основе ПСП, которая генерирует случайные числа в диапазоне от 0 до $L-1$.

В зависимости от численного значения, сформированного ПСП, функция отображения f_{L-1} преобразует сигнал s_i из пространства сигналов S^M в один из сигналов $s_{i,j}$ пространства S_i^N . При этом для значения ПСП равного 0 сигнал s_i преобразуется в сигнал $s_{i,0}$ без изменения своего положения. В то время как для других формируемых ПСП числовых значений, сигнал s_i преобразуется в сигнал $s_{i,j}$ с изменением своего положения в пространстве S_i^N , которое задает функция отображения f_{L-1} .

В общем случае, воздействующая на канал связи аддитивная помеха (АП) приводит к деструктивным последствиям, связанным с изменением положения сигнала в пространстве параметров.

В рассматриваемой ситуации полагаем, что воздействующие АП являются структурными, т.е. их параметры соответствуют пространству параметров передаваемым по КС сигналов.

В результате на вход радиоприемного устройства СРС будет воздействовать аддитивная смесь сигнала и помех, характеризуемая пространством наблюдений $\hat{S}^N = \{\hat{s}_{i,j}\}$. Учитывая, что пространства сигналов S^M и S^N являются взаимно-однозначными, то существует взаимно-однозначное соответствие между элементами S^M и S^N , и можно говорить об отображении пространства сигналов S^N на пространство S^M , $f_{L-1}^{-1} : S^N \rightarrow S^M$, где f_{L-1}^{-1} – обратное отображение [36].

Будем предполагать, что с использованием операции обратного отображения f_{L-1}^{-1} перед решающим устройством демодулятора выполняется преобразование пространства наблюдений $\hat{S}^N = \{\hat{s}_{i,j}\}$ в пространство наблюдений $\hat{S}^M = \{\hat{s}_i\}$ (рис. 2).

Указанное преобразование позволяет при принятии решения в процессе демодуляции использовать пространство сигналов S^M .

Такой подход обеспечит уменьшение ошибок, обусловленных последствием деструктивных воздействий, возникающих на этапе принятия решения.

После демодуляции кодовые символы в декодере канального

кода (ДКК) преобразуются в информационные и поступают к получателю сообщения (ПС).

Представленная с указанных позиций модель СРС позволяет определить порядок синтеза сигналов сложных структур, с одной стороны, обладающих повышенными свойствами структурной сложности [23, 38], а с другой – в наибольшей степени адаптированной к параметрам канала, а также устройствам их формирования и обработки [39-42].

3. Формализация процедур формирования сигналов сложных структур. Реализацию предложенного подхода рассмотрим на примере формирования сигналов квадратурной модуляции, в частности, сигналов квадратурной амплитудной модуляции (КАМ) [43].

Представим конstellационную диаграмму сигналов КАМ- M , где $M = 2^k$ k – четное ($M = 4, 16, 64, \dots$), на комплексной плоскости в виде квадрата со сторонами размером $\sqrt{M} \times \sqrt{M}$. При этом значение \sqrt{M} будем полагать целым натуральным числом.

Осуществим процесс преобразований пространств сигналов при $M = 4$ (сигнал КАМ-4) и $N = 16$ (сигнал КАМ-16). Определим сигнальные точки созвездия КАМ-16 к соответствующим пространствам сигналов $S^M = \{s_i\}$ и $S^N = \{s_{i,j}\}$ таким образом, чтобы соблюдались формализованные свойства пространств сигналов (рис. 3).

Здесь верхние левые сигнальные точки в каждом квадранте КАМ-16 определены к пространству сигналов S^M , представляющему собой сигнал КАМ-4 $S^4 = \{s_0, s_1, s_2, s_3; d_{\min}^{(4)}\}$, и условно обозначены в виде квадратов. Стрелками определены функции отображения $f_{L-1} = \{f_0, f_1, f_2, f_3\}$, которые в зависимости от значения числа, сформированного ПСП, преобразуют сигнальную точку КАМ-4 в одну из сигнальных точек КАМ-16 (обозначены черными кружками).

При формировании ПСП числа 0, функция отображения f_0 никаких перемещений не выполняет, поэтому сигнальные точки КАМ-4 (s_0, s_1, s_2, s_3) имеют одинаковое положение в квадратной решетке с сигнальными точками КАМ-16 ($s_{0,0}, s_{1,0}, s_{2,0}, s_{3,0}$). Таким образом, каждая сигнальная точка s_i КАМ-4 отображается в одну из сигнальных точек $s_{i,L-1}$ пространства сигналов S_i^{16} i -го квадранта КАМ-16.

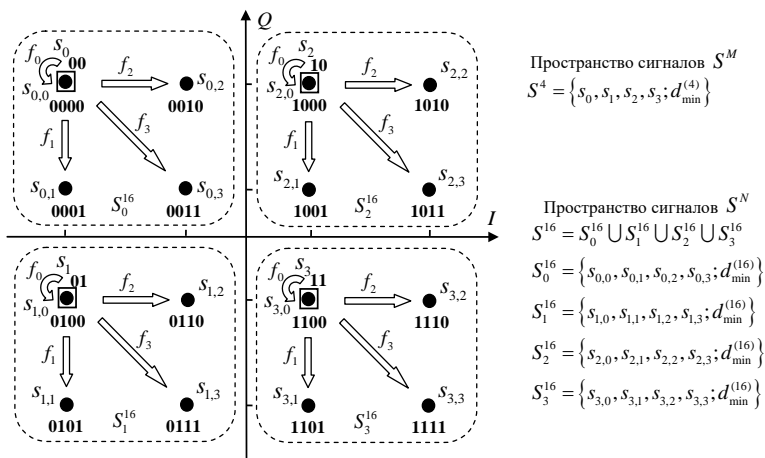


Рис. 3. Преобразование сигнала КАМ-4 в сигнал КАМ-16

Если присвоить сигнальным точкам s_i КАМ-4 кодовые комбинации, то функции отображения f_{L-1} можно задать посредством кодирования или изменения координат сигнальной точки.

Для преобразования сигнала КАМ-4 в сигнал КАМ-16 функции отображения f_{L-1} в зависимости от чисел, формируемых генератором ПСП, обобщим в виде таблицы 1.

Таблица 1. Функции отображения f_{L-1} сигналов s_i КАМ-4 в сигналы $s_{i,L-1}$ КАМ-16

ПСП L	f_{L-1}	Кодирование								Изменение координат	
		$s_0 \rightarrow s_{0,L-1}$		$s_1 \rightarrow s_{1,L-1}$		$s_2 \rightarrow s_{2,L-1}$		$s_3 \rightarrow s_{3,L-1}$		ΔI	ΔQ
0 (00)	f_0	$s_{0,0}$	0000 (0)	$s_{1,0}$	0100 (4)	$s_{2,0}$	1000 (8)	$s_{3,0}$	1100 (12)	0	0
1 (01)	f_1	$s_{0,1}$	0001 (1)	$s_{1,1}$	0101 (5)	$s_{2,1}$	1001 (9)	$s_{3,1}$	1101 (13)	0	$-d_{\min}^{(16)}$
2 (10)	f_2	$s_{0,2}$	0010 (2)	$s_{1,2}$	0110 (6)	$s_{2,2}$	1010 (10)	$s_{3,2}$	1110 (14)	$d_{\min}^{(16)}$	0
3 (11)	f_3	$s_{0,3}$	0011 (3)	$s_{1,3}$	0111 (7)	$s_{2,3}$	1011 (11)	$s_{3,3}$	1111 (15)	$d_{\min}^{(16)}$	$-d_{\min}^{(16)}$

Согласно таблице 1 кодирование выполняется путем добавления к кодовой комбинации, присвоенной сигналу s_i КАМ-4, младших разрядов, представляющих собой двоичную форму числа, сформированного ПСП. В результате такой операции сигналу s_i КАМ-4 присваива-

ется кодовая комбинация сигнала $s_{i,L-1}$ КАМ-16, что эквивалентно изменению положения сигнальной точки s_i КАМ-4 в i -м квадранте квадратной решетки КАМ-16 путем изменения ее координат на значения ΔI и ΔQ , представленные в таблице 1.

Учитывая, что при формировании искомой конструкции используется модулятор сигнала КАМ-16, то функции отображений f_{L-1} реализуются путем кодирования исходного бита.

Таким образом, ПСП не вносит изменения в передаваемую информацию, а изменяет только структуру формируемой конструкции при переходе от КАМ-4 к КАМ-16, что и обеспечивает усложнение структуры формируемого результирующего сигнала.

Поскольку на входе радиоприемного устройства воздействует аддитивная смесь сигнала и помех, характеризуемая в виде пространства наблюдений $\hat{S}^N = \{\hat{s}_{i,j}\}$, то для преобразования ее в пространство $\hat{S}^M = \{\hat{s}_i\}$ необходимо задать обратную функцию отображения f_{L-1}^{-1} . Так как указанные преобразования выполняются до демодулятора, то обратные функции отображения f_{L-1}^{-1} реализуются на основе изменения координат дискретных отсчетов принятой смеси сигнала и помех $\hat{s}_{i,j}$ на значения ΔI и ΔQ , представленные в таблице 1.

Такой подход позволит на приемной стороне использовать демодулятор КАМ-4 с расположением сигнальных точек s_i , представленным на рисунке 3. Обозначим сформированную указанным способом конструкцию как КАМ-4(16), полагая, что прием сигнала осуществляется демодулятором КАМ-4, а структура сигнала на передающей стороне формируется модулятором КАМ-16.

Расположение сигнальных точек сигнала КАМ-4(16) соответствует одному из четырех возможных сигналов КАМ-4, полученных путем разбиения сигнала КАМ-16 по методу Унгербоeka [34]. Тогда можно полагать, что минимальное евклидово расстояние сигнала КАМ-4(16) определяется минимальным евклидовым расстоянием сигнала КАМ-16 следующей зависимостью $d_{\min}^{4(16)} = 2d_{\min}^{16}$. При этом сигнальное созвездие КАМ-4(16) отличается от классического сигнального созвездия КАМ-4 тем, что его сигнальные точки сдвинуты симметрично относительно начала координат.

Процедуру формирования сигналов вида КАМ- $M(N)$, построенных на основе квадратной решетки, определим в виде следующих этапов:

– произвести выбор двумерной конструкции КАМ- N с квадратной решеткой, где $N = 2^n$, $n \geq 4$ – четное число; d_{\min}^N – минимальное евклидово расстояние СК КАМ- N ;

– разбить квадратную решетку исходной конструкции КАМ- N на M непересекающихся пространств по $L = N / M$ сигнальных точек, где $M = 2^k$, $k < n$ – четное число; при этом новые пространства должны быть идентичны, симметричны, а также иметь квадратную форму;

– определить в каждом из M пространств по одной сигнальной точке, которые сформируют СК КАМ- $M(N)$, и присвоить им номера $s_{i,0}$, где i – кодовое слово в соответствии с кодом отображения КАМ- $M(N)$, причем сигнальные точки формируемой СК выбираются таким образом, чтобы средние энергии (E_S) сигналов КАМ- $M(N)$ и КАМ- N были бы равны и выполнялось бы условие $d_{\min}^{M(N)} = \sqrt{L}d_{\min}^N$;

– присвоить с соблюдением геометрической симметрии в каждом формируемом пространстве остальным сигнальным точкам номера $s_{i,j}$, где $j = 0, \dots, L-1$ – число, определяемое ПСП.

4. Исследование свойств сигнальных конструкций КАМ- $M(N)$. Исследование свойств, разработанных сложных сигнальных конструкций КАМ- $M(N)$, проведем на основе показателей, определяющих их помехоустойчивость и структурную сложность.

Анализ помехоустойчивости сигналов КАМ- $M(N)$ осуществим на основе показателей достоверности передачи информации.

В соответствии с [34] достоверность передачи сообщений с использованием сигнальных конструкций с КАМ характеризуется вероятностью битовой ошибки P_b .

Для гауссова канала вероятность появления битовой ошибки при модуляции КАМ- M определяется следующим образом [34, 44]:

$$P_b^M = \frac{2(1 - M^{-1/2})}{\log_2 \sqrt{M}} Q \left(\sqrt{\left(\frac{6 \log_2 \sqrt{M}}{M-1} \right) \frac{E_b}{N_0}} \right), \quad (1)$$

где $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2) dt$ – функция Лапласа; E_b / N_0 – отношение энергии бита к спектральной плотности мощности шума.

При этом минимальное евклидово расстояние для сигналов

КАМ- M на основе квадратной решетки в [36] определено как

$$d_{\min}^M = \sqrt{\frac{6 \log_2 \sqrt{M} E_b}{M-1}} = \sqrt{\frac{6 E_S}{M-1}}, \quad (2)$$

где E_S – средняя энергия сигнала КАМ- M .

С учетом (2) выражение (1) примет следующий вид

$$P_b^M = \frac{2(1-M^{-1/2})}{\log_2 \sqrt{M}} Q\left(\frac{d_{\min}^M}{\sqrt{N_0}}\right). \quad (3)$$

Определение вероятности битовой ошибки сигналов КАМ- M согласно выражениям (2) и (3) справедливо для сигнальных созвездий, у которых центр тяжести ансамбля сигналов совпадает с началом координат. В связи с тем, что сигнальные точки СК КАМ- $M(N)$ сдвинуты относительно начала координат и не соответствуют данному условию, то для определения минимального евклидова расстояния $d_{\min}^{M(N)}$ необходимо использовать условие формирования СК КАМ- $M(N)$: $d_{\min}^{M(N)} = \sqrt{L} \cdot d_{\min}^N$.

При этом сигнальное созвездие КАМ- N , на основе которого формируется СК КАМ- $M(N)$, имеет центр тяжести ансамбля сигналов, совпадающий с началом координат. Поэтому для определения d_{\min}^N можно использовать выражение (2) и тогда получим

$$d_{\min}^{M(N)} = \sqrt{L} \times d_{\min}^N = \sqrt{\frac{N}{M}} \sqrt{\frac{6 E_S}{N-1}} = \sqrt{\frac{6 N E_S}{M(N-1)}}. \quad (4)$$

Учитывая, что средние энергии (E_S) сигналов КАМ- $M(N)$ и КАМ- N равны, применим (4) в выражении (3) и получим вероятность появления битовой ошибки в гауссовом канале для КАМ- $M(N)$ в следующем виде:

$$P_b^{M(N)} = \frac{2(1-M^{-1/2})}{\log_2 \sqrt{M}} Q\left(\sqrt{\left(\frac{6 N \log_2 \sqrt{M}}{M(N-1)}\right) \frac{E_b}{N_0}}\right). \quad (5)$$

Результаты расчета помехоустойчивости СК КАМ-4 и КАМ-

4(16) согласно выражениям (1) и (5) представлены на рисунке 4.

Из анализа результатов, представленных на рисунке 4, следует, что при воздействии на СРС только аддитивного белого гауссова шума применение СК КАМ-4(16) вместо СК КАМ-4 приведет к снижению помехоустойчивости такой СРС.

В частности, при ошибке $P_b^M = 10^{-4}$ снижение в помехоустойчивости достигает порядка 1 дБ.

Проведем оценку предельной величины снижения помехоустойчивости СРС при использовании СК КАМ- $M(N)$ в каналах с гауссовым шумом. В этих целях воспользуемся результатами исследований, представленных в [45], которые обобщим в виде расчета обобщенного показателя W :

$$W = \frac{h_M^2}{h_{M(N)}^2} \Big|_{P_b^{M(N)} = P_b^M},$$

где $h_M^2 = E_b/N_0$ – отношение энергии бита к спектральной плотности мощности шума КАМ- M ; $h_{M(N)}^2 = E_b^{M(N)}/N_0$ – отношение энергии бита к спектральной плотности мощности шума КАМ- $M(N)$.

При условии равенства $P_b^{M(N)} = P_b^M$ в (1) и (5), после преобразований получим

$$W_{[\text{дБ}]} = 10 \lg \left(\frac{M(N-1)}{N(M-1)} \right). \quad (6)$$

Анализ сложности структуры СК КАМ- $M(N)$ проведем на основе элементов теории, представленной в работе [7]. Сложность структуры определим числом двоичных измерений (ДИЗ), которые необходимо провести для полной формализации структуры СК.

Общее выражение для определения данного показателя примем в виде:

$$S = \log_2 A,$$

где A – допустимые значения произвольного параметра сигнала. Таким параметром может быть выбрана амплитуда, вид модуляции, структура СК и др.

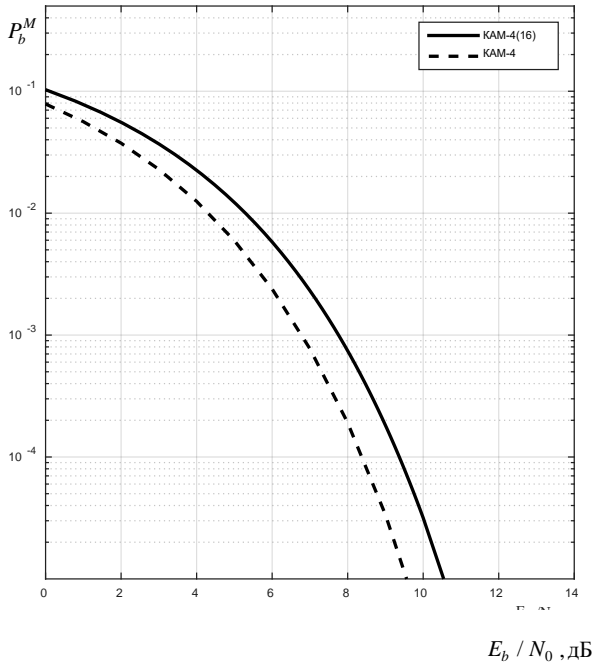


Рис. 4. Помехоустойчивость СК КАМ-4 и КАМ-4(16)

Для СК КАМ- M количество вариантов соответствия каждой точке СК символов с учетом кода Грея определяется размерностью M , поэтому $S = \log_2 M$. Для полной формализации структуры СК КАМ- $M(N)$ необходимо дополнительно учесть варианты соответствия функции отображения f_{L-1} сигнальным точкам квадранта квадратной решетки СК (рис. 3) и числам, формируемым ПСП, что равнозначно количеству перестановок из L и может быть определено как $L!$.

Тогда сложность структуры СК КАМ- $M(N)$ определим в следующем виде

$$S = \log_2 (M(N/M)!). \quad (7)$$

Результаты расчета выбранных параметров СК КАМ- $M(N)$ представлены в таблице 2. Из анализа полученных результатов следует, что с увеличением L сложность структуры СК КАМ- $M(N)$ возрастает.

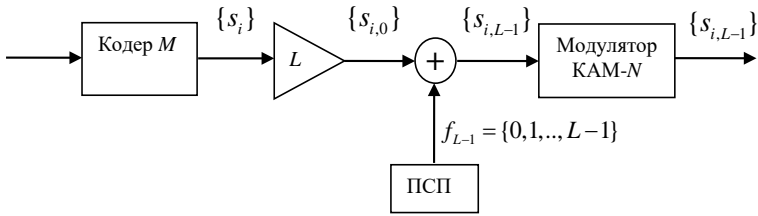
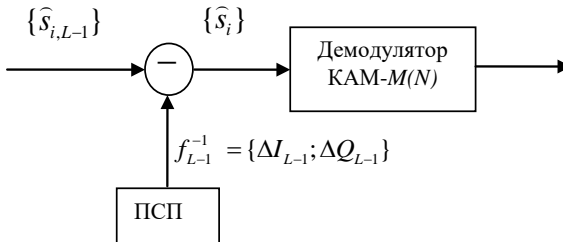
ет, однако при этом потери в помехоустойчивости увеличиваются незначительно. Так, например, для СК КАМ-4(16) и КАМ-4(256) сложность структуры увеличилась в 42 раза, а потери в помехоустойчивости составляют не более 0,26 дБ.

Таблица 2. Результаты расчета параметров СК КАМ- $M(N)$

КАМ- M	S , ДИЗ	КАМ- $M(N)$	$L=N/M$	W , дБ	S , ДИЗ
КАМ-4	2	КАМ-4(16)	4	0,97	7
		КАМ-4(64)	16	1,18	47
		КАМ-4(256)	64	1,23	298
КАМ-16	4	КАМ-16(64)	4	0,21	9
		КАМ-16(256)	16	0,26	49
КАМ-64	6	КАМ-64(256)	4	0,05	11

Вместе с тем, реализация разработанных алгоритмов наиболее целесообразна для повышения энергетической скрытности реализованной СРС на основе методов доступа абонентов к каналам радиосвязи. Действительно, нелегитимным абонентам неизвестен способ формирования и обработки сигналов, что определяет возможность доступа в СРС только тем абонентам, которым выделены ключи ПСП и функции преобразования сигналов. Причем разработанные подходы к формированию и обработки сигналов, в отличие от известных [18, 22], не предъявляют высоких требований к параметрам чувствительности радиоприемных устройств СРС.

5. Результаты эксперимента по моделированию СРС с КАМ- $M(N)$. Для проверки адекватности полученных теоретических результатов проведено моделирование СРС, использующей разработанный способ формирования СК. Для этого разработаны структурные схемы модулятора и демодулятора СК КАМ- $M(N)$, представленные на рисунках 5 и 6 соответственно. В схеме модулятора кодер M разбивает двоичный цифровой поток на блоки длиной $k = \sqrt{M}$ бит и переводит их в десятичную форму счисления, формируя этим пространство сигналов КАМ- M ($S^M = \{s_i\}$). Далее с помощью умножения на $L = N / M$ и прибавления десятичного числа, сформированного генератором ПСП $\{0, 1, \dots, L-1\}$, сигналу присваивается номер, который определяет его положение в сигнальном созвездии модулятора КАМ- N ($S^N = \{s_{i,L-1}\}$).

Рис. 5. Структурная схема модулятора СК КАМ- $M(N)$ Рис. 6. Структурная схема демодулятора СК КАМ- $M(N)$

На вход приемника поступают дискретные отсчеты аддитивной смеси сигнала и гауссова шума, образуя пространство наблюдений $\{\widehat{s}_{i,L-1}\}$. Генератор ПСП приемника синхронно с аналогичным генератором ПСП передатчика формирует соответствующие значения координат ΔI и ΔQ , которые поступают на вычитающее устройство. За счет изменения координат дискретных отсчетов пространство наблюдений $\{\widehat{s}_{i,L-1}\}$ преобразуется в пространство наблюдений $\{\widehat{s}_i\}$, что позволит для принятия решений использовать демодулятор КАМ- $M(N)$.

Программная реализация системы связи в программной среде Simulink представлена на рисунке 7.

На рисунке 8 представлен пример сигнальных созвездий для модулятора КАМ-16 и для демодулятора СК КАМ-4(16).

С использованием представленной модели в программной среде Simulink проведен эксперимент методом статистических испытаний для оценки помехоустойчивости СК КАМ-4(16), КАМ-4(64), КАМ-

16(64). Результаты приведены на рисунках 9-12.

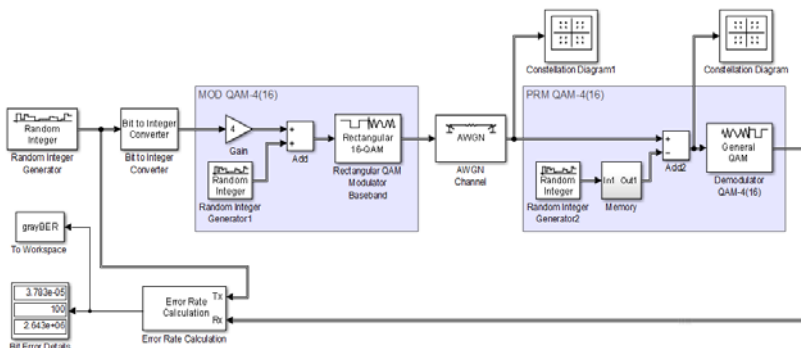


Рис. 7. Модель системы связи в программе Simulink

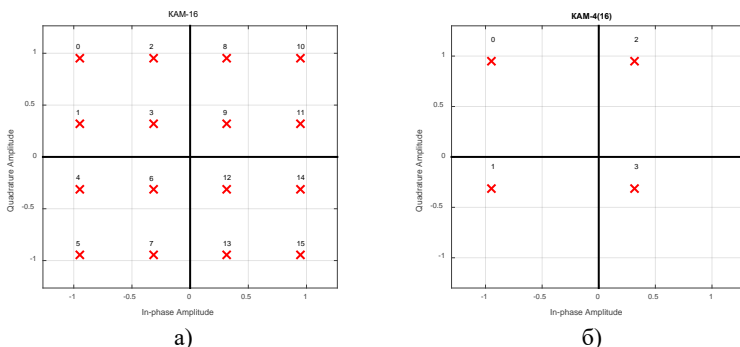
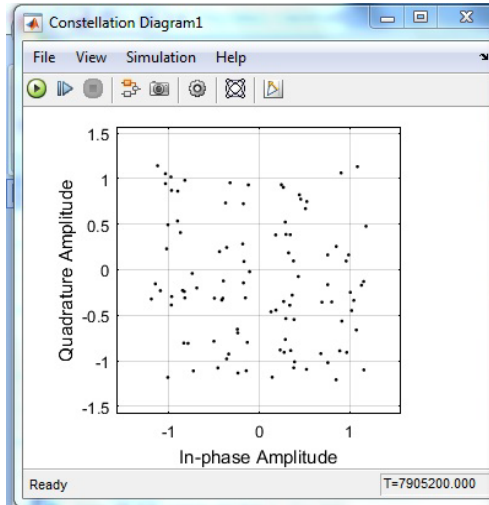
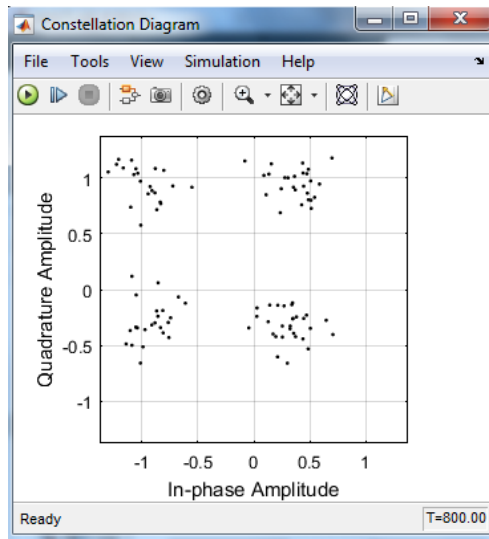


Рис. 8. Сигнальные созвездия: а) модулятора QAM-16; б) демодулятора QAM-4(16)

На рисунке 9 представлены фрагменты констеляционных диаграмм аддитивной смеси гауссова шума и сигнала, действующей на входе приемника (а) и на входе демодулятора QAM-4(16) (б) после проведения преобразований пространств наблюдений $\hat{s}_{i,L-1} \rightarrow \hat{s}_i$ при отношении энергии бита к спектральной плотности мощности шума $E_b / N_0 = 10$ дБ.



а)



б)

Рис. 9. Констеляционная диаграмма: а) пространство наблюдений $\{\hat{s}_{i,L-1}\}$ на входе приемника; б) пространство наблюдений $\{\hat{s}_i\}$ на входе демодулятора КАМ-4(16)

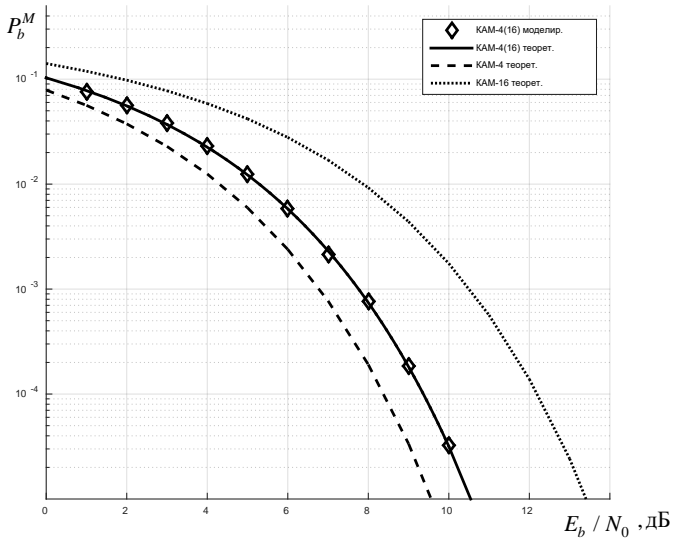


Рис. 10. Помехоустойчивость СК КАМ-4(16)

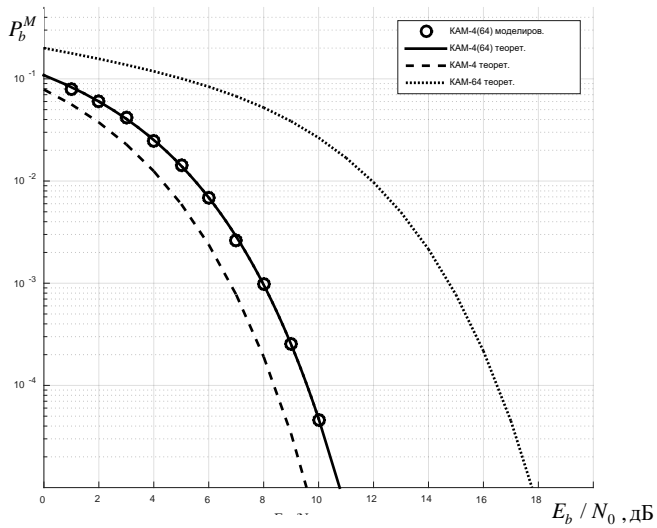


Рис. 11. Помехоустойчивость СК КАМ-4(64)

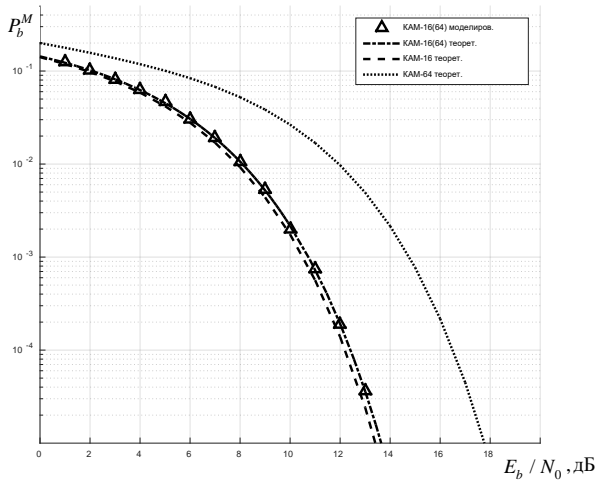


Рис. 12. Помехоустойчивость СК КАМ-16(64)

Анализ результатов, представленных на рисунках 10-12, позволяет сделать следующее заключение. Теоретические выводы полностью подтверждаются данными компьютерного моделирования. Так, графики зависимости вероятности битовой ошибки от отношения энергии бита к спектральной плотности мощности шума, характеризующие помехоустойчивость СК сложной структуры, полученные путем моделирования и построенные в соответствии с выражением (5) полностью совпадают. Таким образом, можно утверждать, что разработанные теоретические основы формирования и обработки СК сложной структуры являются правомерными и обоснованными.

6. Заключение. Разработанные подходы к формированию сигнальных конструкций сложной структуры в своей совокупности обеспечивают повышение сложности структуры при незначительном снижении помехоустойчивости. При этом нелегитимным абонентам ограничивается доступ к формируемым СРС каналам за счет изменения структуры излучаемого радиосигнала, что позволяет отнести разработанные методы к классу СРС на основе сигналов большой базы. Отличительной особенностью предложенного подхода формирования сигналов сложной структуры является то, что при его использовании не задействуется дополнительно частотный ресурс СРС, поскольку увеличение базы сигналов достигается на основе расширения пространства сигналов.

Применение сигналов сложной структуры на основе сигналов с

квадратурной амплитудной модуляцией позволит за счет расширения пространства сигналов КАМ-4 в 4 раза увеличить сложность структуры в 3,5 раза, а при увеличении пространства сигналов в 16 раз сложность структуры увеличивается в 23,5 раза, при этом помехоустойчивость снизится на 1,18 дБ. При расширении пространства сигналов КАМ-16 в 4 раза сложность возрастает в 2,25 раза, но при этом помехоустойчивость снижается только на 0,21 дБ.

Таким образом, в статье определяется класс задач по разработке теоретических положений, формализующих свойства СРС с сигналами сложных структур на основе расширения пространства сигналов.

Литература

1. Мальцев Г.Н., Харченко А.В., Штанько С.В. Сравнительный анализ методов ограничения и управления доступом в радиотехнических системах с широковещательной передачей информации // Труды Военно-космической академии имени А.Ф. Можайского. 2018. № 665. С. 56-68.
2. Макаренко С.И. Информационное противоборство и радиоэлектронная борьба в сетевых войнах начала XXI века // Санкт-Петербург. 2017.
3. Штанько С.В. Ограничение несанкционированного доступа в радиотехнических системах с широковещательной передачей информации // Информационно-управляющие системы. 2018. № 5 (96). С. 57-65.
4. Moldovyan D.N., Moldovyan N.A., Moldovyan A.A. Commutative encryption method based on Hidden logarithm problem. Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software. 2020. vol. 13. no. 2. pp. 54-68.
5. Moldovyan A.A., Phieu N.H., Nguyen H.M., Tran C.M. Digital signature algorithms based on Hidden discrete logarithm problem. Advances in Intelligent Systems and Computing. 2020. vol. 1014. pp. 1-12.
6. Kuzmin A.S., Markov V.T., Mikhalev A.A., Mikhalev A.V., Nechaev A.A. Cryptographicalgorithms on groups and algebras. Journal of Mathematical Sciences. 2017. vol. 223. no. 5. pp. 629-641.
7. Каневский З.М., Литвиненко В.П. Теория скрытности // Воронеж: Изд-во Воронеж. гос. ун-та. 1991. 142 с.
8. Пантенков Д.Г., Литвиненко В.П. Алгоритмы формирования и обработки радиосигналов командно-телеметрической радиолинии и технические предложения по их реализации // Вестник Воронежского государственного технического университета. 2020. Т. 16. № 2. С. 90-105.
9. Гавришев А.А. Применение КС-энтропии для количественного анализа систем связи гражданской обороны // Сибирский пожарно-спасательный вестник. 2020. № 3 (18). С. 47-54.
10. Сизов А.С., Цыплаков Ю.В. Анализ сложных сигналов, применяемых в радиолокации и скрытной радиосвязи // Телекоммуникации. 2020. № 5. С. 33-40.
11. Zhilyakov E.G., Belov S.P., Ursol D.V. About signals allowing to provide sustainability to impacts of short-term and or focused on spectrum interference. International Journal of Engineering and Technology (UAE). 2018. vol. 7. no. 2. pp. 16-20.
12. Киченко А.Н., Савельев А.В., Шишаков К.В. Анализ сигнально-кодовых конструкций в задачах повышения энергетической скрытности радиолиний // Интеллектуальные системы в производстве. 2020. Т. 18. № 4. С. 11-20.

13. Kirillov S.N., Lisnichuk A.A., Lukasin I.V., Pokrovskij P.S. Methods to form both anti-interference and hiding radio signals for prospective communication systems. International Siberian Conference on Control and Communications. (SIBCON'2019) Proceedings. 2019. P. 8729649.
14. Дворников С.В., Манаенко С.С., Дворников С.С. Параметрическая мимикрия сигналов, модулированных колебаниями и сформированных в различных функциональных базисах // Информационные технологии. 2015. Т. 21. № 4. С. 259-263.
15. Kirillov S.N., Lisnichuk A.A., Lukashin I.V., Pokrovskij P.S. Algorithms to form and process signals for multilevel defense of transmitted data based on the principles of cognitive and software-defined radio. Moscow Workshop on Electronic and Networking Technologies. (MWENT'2018) Proceedings. 2018. pp. 1-6.
16. Финк Л.М. Теория передачи дискретных сообщений // М.: Сов. Радио. 1970. 718 с.
17. Миддлтон Д. Введение в статистическую теорию связи. Том 1. / Пер. с англ. Б. А. Смиренина // Под ред. и с предисл. Б.П. Левина. М.: Сов. Радио. 1961. 782 с.
18. Феер К. Беспроводная цифровая связь. Методы модуляции и расширения спектра. Пер. с англ. / Под ред. В. И. Журавлева // М.: Радио и связь. 2000. 520 с.
19. Варакин Л.Е. Теория сложных сигналов // М.: Связь. 1970. 317 с.
20. Pickholtz R.L., Schilling D.L., Milstein L.B. Theory of Spread-Spectrum Communications – A Tutorial. IEEE Trans. Commun. 1982. vol. 30 (5). pp. 855–884.
21. Anderson J.B. Called Modulation Systems. N.Y.: Kluwer Academic Publishers. 2003. P. 544.
22. Дворников С.В., Пшеничников А.В., Манаенко С.С. Модель фазоманипулированного широкополосного сигнала с программной перестройкой рабочей частоты // Телекоммуникации. 2017. № 9. С. 8-12.
23. Борисов В.И., Зинчук В.М., Лимарев А.Е. Помехозащищенность систем радиосвязи расширением спектра сигналов методом псевдослучайной перестройки рабочей частоты / под ред. В. И. Борисова; изд. 2-е, перераб. и доп. // М.: РадиоСофт. 2008. 512 с.
24. Беккиев А.Ю., Борисов В.И. Оценка помехозащищенности каналов радиосвязи в условиях действия помех от средств радиоэлектронной борьбы // Радиотехника и электроника. 2019. Т. 64. № 9. С. 891-901.
25. Дворников С.В., Пшеничников А.В., Русин А.А. Обобщенная функциональная модель радиолинии с управлением её частотным ресурсом // Вопросы радиоэлектроники. Серия: Техника телевидения. 2016. № 3. С. 49-56.
26. Дворников С.В., Дворников С.С., Пшеничников А.В. Аппарат анализа частотного ресурса для режима псевдослучайной перестройки рабочей частоты // Информационно-управляющие системы. 2019. № 4 (101). С. 62-68.
27. Пшеничников А.В. Оценка статистических параметров рабочих частот для модели радиолинии в конфликтной ситуации // T-Comm: Телекоммуникации и транспорт. 2017. Т. 11. № 1. С. 4-9.
28. Тузов Г.И. Статистическая теория приема сложных сигналов // М.: Сов. Радио. 1977. 116 с.
29. Дворников С.В., Пшеничников А.В. Формирование спектрально-эффективных сигнальных конструкций в радиоканалах передачи данных контрольно-измерительных комплексов // Известия высших учебных заведений. Приборостроение. 2017. Т. 60. № 3. С. 221-228.
30. Прокис Дж. Цифровая связь / Пер. с англ.; Под ред. Д.Д. Кловского // М.: Радио и связь. 2000. 800 с.
31. Кловский Д.Д. Передача дискретных сообщений по радиоканалам // М.: Радио и связь. 1982. 304 с.

32. Дворников С.В., Пшеничников А.В. Помехозащищенная модель радиолинии в условиях динамического преднамеренного воздействия // Известия высших учебных заведений России. Радиоэлектроника. 2017. № 2. С. 16-22.
33. Месарович М., Такахара Я. Общая теория систем. Математические основы / Пер. с англ. Э.Л. Наппельбаума // Под редакцией С.В. Емельянова. М.: Издательство «Ширь». 1978. 315 с.
34. Скляр Б. Цифровая связь: теоретические основы и практическое применение / Пер. с англ. Е.Е. Грозы и др. 2-е изд. М.: Вильямс, 2016. 1099 с.
35. Хелстром К. Статистическая теория обнаружения сигналов // М.: Издательство иностранной литературы. 1963. 432 с.
36. Wilson S.G. Digital Modulation and Coding. Prentice Hall. 1996. P. 667.
37. Френкс Л. Теория сигналов / Пер. с англ. // М.: Сов.Радио. 1974. 344 с.
38. Дворников С.В., Пшеничников А.В., Манаенко С.С., Дворников С.С. Метод формирования многопозиционных помехозащищенных сигнальных конструкций // Информационные технологии. 2017. Т. 23. № 9. С. 669-676.
39. Mehrota A. Analog and Digital Systems. Artech House. London. 1996. P. 205.
40. Ungerboeck G. Trellis-coded modulation with redundant signal sets Part II: State of the art. Communications Magazine. IEEE. 1987. Vol. 25. Issue: 2. pp. 5-11.
41. Torrieri D.J. Principles of secure communication systems. Dedham. MA.: Artech House Inc. 1985. P. 286.
42. Пшеничников А.В. Модель многопозиционной помехозащищенной сигнальной конструкции на основе QPSK модуляции // Информация и космос. 2017. № 2. С. 48-52.
43. Cahn C.R. Combined Digital Phase and Amplitude Modulation Communication Systems. IRE Transaction on Communications Systems. Vol. CS-8. 1960. pp. 150-154.
44. Korn I. Digital Communications. N. Y.: Van Nostrand Reinhold Co. 1985. P. 676. p.
45. Буга Н.Н. Информация и электрические сигналы // Л.: ЛКВВИА. 1960. 71 с.

Манаенко Сергей Сергеевич — канд. техн. наук, старший преподаватель, кафедра радиосвязи, Военная академия связи. Область научных интересов: построение помехозащищенных систем радиосвязи, способов формирования и обработки сигналов сложных структур. Число научных публикаций — 78. manaenkoss@mail.ru; Тихорецкий проспект, 3, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-98-11.

Дворников Сергей Викторович — д-р техн. наук, профессор, профессор кафедры, кафедра радиотехнических и оптоэлектронных комплексов, Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет аэрокосмического приборостроения». Область научных интересов: построение помехозащищенных систем радиосвязи, способов формирования и обработки сигналов сложных структур. Число научных публикаций — 283. practicdsv@yandex.ru; Тихорецкий проспект, 3, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-98-11.

Пшеничников Александр Викторович — д-р техн. наук, доцент, заместитель начальника кафедры, кафедра радиосвязи, Военная академия связи. Область научных интересов: построение помехозащищенных систем радиосвязи, способов формирования и обработки сигналов сложных структур. Число научных публикаций — 147. sigacooz77@mail.ru; Тихорецкий пр., 3, 194064, Санкт-Петербург, Россия; р.т.: +7(812)247-98-11.

S. MANAENKO, S. DVORNIKOV, A. PSHENICHNIKOV
**THEORETICAL ASPECTS IN FORMING COMPLEX STRUCTURE
SIGNAL**

Manaenko S., Dvornikov S., Pshenichnikov A. Theoretical Aspects in Forming Complex Structure Signal.

Abstract. Receiving and transmitting paths of modern radio communication systems are built on the basis of an open structure that provides hierarchical differentiation of access to the provided telecommunication services. However, this approach does not exclude the possibility of access to the transmitted content by unauthorized users. Hiding information by methods of cryptographic protection in such a situation only activates additional interest in transmission, therefore the most pragmatic solution is to use signals of a complex structure, which significantly complicate or even exclude the extraction of information from them by third-party users. The problem of regulating access selection in the development and design radio system elements is rather multifaceted and has a high degree of complexity. One of the directions for solving problems in this subject area is based on the well-known approaches to expanding the signal base, however, algorithms for their practical implementation were obtained without taking into account the limitations on the allocated resource and the very fact of using these algorithms. Based on the theory of systems and the general theory of communication, an approach to the formation of signal structures of a complex structure has been developed, which ensures an increase in the properties of their structural secrecy in relation to unauthorized users. At the same time, the known solutions at the physical level of signal spaces were refined, which made it possible to formalize the procedures for the formation of radio signals with specified properties. The method of formalizing the function of displaying the signal space based on the allocation of stochastic properties of pseudo-random sequences has been substantiated, which made it possible to ensure the uncertainty of their structure in case of unauthorized processing. The approbation of the proposed approach is given on the example of the formation of quadrature modulation signals, taking into account the subsequent analysis of their properties from various positions of legitimate and illegitimate users. The results obtained confirm the uncertainty during illegitimate processing with a slight deterioration in the noise immunity properties of radio communication systems. In general, this allows to conclude the adequacy of theoretical solutions. As an example, constellation diagrams of signals at the output of a quadrature receiver are presented. The set of proposed technical solutions presented in the work determines the novelty of this approach. The scientific problem to be solved belongs to the class of problems of synthesis of signals of complex structures.

Keywords: signals of complex structure, structural secrecy of signals, energy secrecy of signals, synthesis of signals of complex structure, signal structures.

Manaenko Sergey — Ph.D., Senior lecturer, Department of radio communication, Military academy of communications. Research interests: construction of noise-immune radio communication systems, methods of generating and processing signals of complex structures. The number of publications — 78. manaenkoss@mail.ru; 3, Tikhoretsky pr., 194064, St. Petersburg, Russia; office phone: +7(812)247-98-11.

Dvornikov Sergey — Ph.D., Dr.Sci., Professor, Professor of the department, Department of radio engineering and optoelectronic complexes, Federal State Autonomous Educational Institution of Higher Education "St. Petersburg State University of Aerospace Instrumentation". Research interests: construction of noise-immune radio communication systems, methods of generating and processing signals of complex structures. The number of

publications — 283. practicsv@yandex.ru; 3, Tikhoretsky pr., 194064, St. Petersburg, Russia; office phone: +7(812)247-98-11.

Pshenichnikov Alexander — Ph.D., Dr.Sci., Associate Professor, Deputy head of department, Department of radio communication, Military academy of communications. Research interests: construction of noise-immune radio communication systems, methods of generating and processing signals of complex structures. The number of publications — 147. siracooz77@mail.ru; 3, Tikhoretsky pr., 194064, St. Petersburg, Russia; office phone: +7(812)247-98-11.

References

1. Mal'cev G.N., Harchenko A.V., Shtan'ko S.V. [Comparative analysis of methods of restricting and controlling access in radio engineering systems with broadcasting information] *Trudy Voenno-kosmicheskoy akademii imeni A.F. Mozhajskogo. [Proceedings of the Mozhaisky Military Space Academy].* 2018. no. 665. pp. 56-68. (In Russ.).
2. Makarenko S.I. *Informacionnoe protivoborstvo i radioelektronnaja bor'ba v setecentricheskix vojnax nachala XXI veka [Information confrontation and electronic warfare in network-centric wars at the beginning of the XXI century].* Sankt-Peterburg. 2017. (In Russ.).
3. Shtan'ko S.V. [Restriction of unauthorized access in radio engineering systems with broadcasting information]. *Informacionno-upravljajushhie sistemy. [Information and control systems].* 2018. no. 5 (96). pp. 57-65. (In Russ.).
4. Moldovyan D.N., Moldovyan N.A., Moldovyan A.A. Commutative encryption method based on Hidden logarithm problem. *Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software.* 2020. vol. 13. no. 2. pp. 54-68.
5. Moldovyan A.A., Moldovyan N.A., Phieu N.H., Nguyen H.M., Tran C.M. Digital signature algorithms based on Hidden discrete logarithm problem. *Advances in Intelligent Systems and Computing.* 2020. vol. 1014. pp. 1-12.
6. Kuzmin A.S., Markov V.T., Mikhalev A.A., Mikhalev A.V., Nechaev A.A. Cryptographicalgorithms on groups and algebras. *Journal of Mathematical Sciences.* 2017. vol. 223. no. 5. pp. 629–641.
7. Kanevskij Z.M., Litvinenko V.P. *Teorija skrytnosti [Theory stealth].* Voronezh: Izdvo Voronezh.gos. un-ta. [Voronezh: Publishing Voronezh State. University]. 1991. P. 142. (In Russ.).
8. Pantenkov D.G., Litvinenko V.P. [Algorithms of formation and processing of radio signals of command and telemetry radio lines and technical proposals for their implementation]. *Vestnik Voronezhskogo gosudarstvennogo tehničeskogo universiteta. [Bulletin of voronezh state technical university].* 2020. vol. 16. no. 2. pp. 90-105. (In Russ.).
9. Gavrishiev A.A. [Application of ks-entropy for quantitative analysis of civil defense communication systems]. *Sibirskij požarno-spasatel'nyj vestnik. [Siberian fire and rescue bulletin].* 2020. no. 3 (18). pp. 47-54. (In Russ.).
10. Sizov A.S., Cyplakov Ju.V. [Analysis of composite signals in radiolocation and secured radio communication]. *Telekommunikacii [Telecommunications].* 2020. no. 5. pp. 33-40. (In Russ.).
11. Zhilyakov E.G., Belov S.P., Ursol D.V. About signals allowing to provide sustainability to impacts of short-term and or focused on spectrum interference. *International Journal of Engineering and Technology (UAE).* 2018. vol. 7. no. 2. pp. 16-20.

12. Kichenko A.N., Savel'ev A.V., Shishakov K.V. [Analysis of signal-code structures in the tasks of increasing the energy stealth of radios]. *Intellektual'nye sistemy v proizvodstve*. [Intelligent systems in production]. 2020. vol. 18. no. 4. pp. 11-20. (In Russ.).
13. Kirillov S.N., Lisnichuk A.A., Lukasin I.V., Pokrovskij P.S. Methods to form both anti-interference and hiding radio signals for prospective communication systems. *International Siberian Conference on Control and Communications*. (SIBCON'2019) Proceedings. 2019. P. 8729649.
14. Dvornikov S.V., Manaenko S.S., Dvornikov S.S. [Parametric mimicry signals modulated oscillations formed in various functional basis]. *Informacionnye tehnologii*. [Information technology]. 2015. vol. 21. no. 4. pp. 259-263. (In Russ.).
15. Kirillov S.N., Lisnichuk A.A., Lukashin I.V., Pokrovskij P.S. Algorithms to form and process signals for multilevel defense of transmitted data based on the principles of cognitive and software-defined radio. *Moscow Workshop on Electronic and Networking Technologies*. (MWENT'2018) Proceedings. 2018. pp. 1-6.
16. Fink L.M. *Teorija peredachi diskretnyh soobshhenij* [Discrete message transmission theory]. M.: Sov. Radio. 1970. P. 718. (In Russ.).
17. Middleton D. *An introduction to statistical communication theory*. Volume 1. 1960. Russ. ed.: Middleton D. *Vvedenie v statisticheskuyu teoriju svyazi*. Tom 1. M.: Sov. Radio. 1961. P. 782.
18. Feher K. *Wireless digital communications. Modulation and spread spectrum applications*. Prentice Hall PTR. 1995. Russ. ed.: Feher K. *Besprovodnaja cifrovaja svjaz'. Metody moduljacji i rasshirenija spectra*. M.: Radio i svjaz'. 2000. P. 520.
19. Varakin L.E. *Teorija slozhnyh signalov*. [Complex signal theory]. M.: Svjaz'. 1970. P. 317. (In Russ.).
20. Pickholtz R.L., Schilling D.L., Milstein L.B. *Theory of Spread-Spectrum Communications – A Tutorial*. *IEEE Trans. Commun.* 1982. vol. 30 (5). pp. 855–884.
21. Anderson J.B. *Called Modulation Systems*. N.Y.: Kluwer Academic Publishers. 2003. P. 544.
22. Dvornikov S.V., Pshenichnikov A.V., Manaenko S.S. [Model of phase-shift keying broadband signal with program operating frequency tuning]. *Telekommunikacii*. [Telecommunications]. 2017. no. 9. pp. 8-12. (In Russ.).
23. Borisov V.I., Zinchuk V.M., Limarev A.E. *Pomehozashhishennost' sistem radiosvyazi rasshireniem spektra signalov metodom psevdosluchajnoj perestrojki rabochej chastoty*. [Interference immunity of radio communication systems by spreading the spectrum of signals by the method of pseudo-random tuning of the operating frequency]. M.: RadioSoft. 2008. P. 512. (In Russ.).
24. Bekkiev A.Ju., Borisov V.I. [Estimation of the anti-interference ability of radio communication channels under electronic warfare conditions]. *Radiotekhnika i jelektronika*. [Journal of Communications Technology and Electronics]. 2019. vol. 64. no. 9. pp. 891-901. (In Russ.).
25. Dvornikov S.V., Pshenichnikov A.V., Rusin A.A. [Generalized functional model of a radio link control its frequency resource]. *Voprosy radiojelektroniki. Serija: Tehnika televidenija*. [Questions of radio electronics. Series: television technology]. 2016. no. 3. pp. 49-56. (In Russ.).
26. Dvornikov S.V., Dvornikov S.S., Pshenichnikov A.V. [Analysis of frequency resource for fhss mode]. *Informacionno-upravljajushhie sistemy*. [Information and Control Systems]. 2019. no. 4 (101). pp. 62-68. (In Russ.).
27. Pshenichnikov A.V. [Estimation of statistical parameters of operating frequencies for radio model in conflict situations]. *T-Comm: Telekommunikacii i transport*. [T-Comm: Telecommunications and transport]. 2017. vol. 11. no. 1. pp. 4-9. (In Russ.).

28. Tuzov G.I. Statisticheskaja teorija priema slozhnyh signalov [Statistical theory of complex signal reception]. M.: Sov. Radio. 1977. P. 116. (In Russ.).
29. Dvornikov S.V., Pshenichnikov A.V. [Design of spectral-efficient signal construction in radio data transmission control and measuring complexes]. Izvestija vysshih uchebnyh zavedenij. Priborostroenie. [Journal of Instrument Engineering]. 2017. vol. 60. no. 3. pp. 221-228. (In Russ.).
30. Proakis J.G. Digital communication. Mc Graw-Hill. 1995. Russ. ed.: Prokis Dzh. Cifrovaja svjaz'. M.: Radio i svjaz'. 2000. P. 800.
31. Klovkij D.D. Peredacha diskretnyh soobshhenij po radiokanalam. [Transmission of discrete messages over radio channels]. M.: Radio i svjaz'. 1982. P. 304. (In Russ.).
32. Dvornikov S.V., Pshenichnikov A.V. [Noise immunity radio link model in dynamic intentional exposure]. Izvestija vysshih uchebnyh zavedenij Rossii. Radiojelektronika. [Izvestia of the Higher Educational Institutions of Russia. Radioelectronics]. 2017. no. 2. pp. 16-22. (In Russ.).
33. Mesarovich M., Takahara Ja. Obshhaja teorija sistem. Matematicheskie osnovy. [General systems theory. Mathematical foundations]. M.: Izdatel'stvo «Shir». 1978. P. 315. (In Russ.).
34. Skljjar B. Cifrovaja svjaz': teoreticheskie osnovy i prakticheskoe primenenie. [Digital Communications: Fundamentals and Applications]. M.: Vil'yams. 2016. P. 1099. (In Russ.).
35. Helstrom K. Statisticheskaja teorija obnaruzhenija signalov. [Statistical signal detection theory]. M.: Izdatel'stvo inostranoj literatury. 1963. P. 432. (In Russ.).
36. Wilson S.G. Digital Modulation and Coding. Prentice Hall. 1996. P. 667.
37. Franks L.E. Signal theory. Prentice Hall Inc Englewood Cliffs. N.J. 1969. Russ. ed.: Frenks L. Teorija signalov. M.: Sov. Radio. 1974. P. 344.
38. Dvornikov S.V., Pshenichnikov A.V., Manaenko S.S., Dvornikov S.S. [A method of forming multiposition jamproof signal constructions]. Informacionnye tehnologii. [Information technology]. 2017. vol. 23. no. 9. pp. 669-676. (In Russ.).
39. Mehrota A. Analog and Digital Systems. Artech House. London. 1996. P. 205.
40. Ungerboeck G. Trellis-coded modulation with redundant signal sets Part II: State of the art. Communications Magazine. IEEE. 1987. Vol. 25. Is. 2. pp. 5-11.
41. Torrieri D.J. Principles of secure communication systems. Dedham. MA.: Artech House Inc. 1985. P. 286.
42. Pshenichnikov A.V. [Model of multi-position noise-proof signal structure based on QPSK modulation]. Informacija i kosmos. [Information and Space]. 2017. no. 2. pp. 48-52. (In Russ.).
43. Cahn C.R. Combined Digital Phase and Amplitude Modulation Communication Systems. IRE Transactions on Communications Systems. Vol. CS-8. 1960. pp. 150-154.
44. Korn I. Digital Communications. N. Y.: Van Nostrand Reinhold Co. 1985. P. 676.
45. Buga N.N. Informacija i jelektricheskie signaly [Informations and electrical signals]. L.: LKVVIA. 1960. P. 71. (In Russ.).

Б.Г. МАЙОРОВ
**СВОЙСТВА ГАРМОНИЧЕСКИХ И СОСТАВНЫХ ПОЛУВОЛН,
ОПРЕДЕЛЕНИЕ ИНТЕРВАЛА РАВНОМЕРНОЙ
ДИСКРЕТИЗАЦИИ ВРЕМЕНИ ЦИФРОВЫХ СИГНАЛЬНЫХ
ПРОЦЕССОРОВ**

Майоров Б.Г. Свойства гармонических и составных полуволи, определение интервала равномерной дискретизации времени цифровых сигнальных процессоров.

Аннотация. При построении автономных систем реального времени (СРВ) необходимо решать задачу оптимальной многозадачной загрузки ряда параллельно функционирующих цифровых сигнальных процессоров. Одним из резервов достижения необходимого результата выступает реализация выборок из сигналов датчиков информации о величине сигнала наиболее редко во времени. При этом необходимо обеспечивать линейную или ступенчатую аппроксимацию сигнала по выборкам с допустимой погрешностью восстановления. Одной из системных задач этих процессоров является фильтрация сигналов или ограничение спектра до частоты среза. Отличительной особенностью предлагаемого в статье подхода является выполнение условия: если измерение этой частоты затруднено (например, в электромеханических средствах СРВ), то для таких сигналов предложено согласовывать максимальные величины параметров гармонической полуволи: погрешность аппроксимации, скорость и ускорение. Исследование открывает перспективу применения новых подходов по дискретизации времени сигналов в амплитудно-временной области и определение для таких сигналов эквивалентной частоты среза спектра сигнала. В настоящей статье получены зависимости величины единицы системного времени ввода-вывода данных от степени согласования между собой максимальных величин параметров сигнала. Математическая модель экстремального поведения сигнала между двумя соседними выборками задана в виде гармонической полуволи. Исследование распространено также на выпуклые составные гармонические функции, по которым сигнал может отклоняться от результатов линейной или ступенчатой аппроксимации сигнала по этим выборкам. Проведено сравнение моделей по величине относительных интервалов дискретизации времени, зависящих от степени согласования максимальных параметров гармонической полуволи. При сравнении, кроме этих максимальных параметров, учтена связь максимальной скорости сигнала с погрешностью аппроксимации выборок ступеньками и связь максимального ускорения сигнала с максимальной погрешностью линейной аппроксимации. Результаты позволяют определять длительности интервалов равномерной дискретизации времени сигнала по результатам обследования объекта управления, обосновывают существенное увеличение интервала дискретизации времени или аналогичное увеличение числа решаемых задач в единицу системного времени.

Ключевые слова: системы управления реального времени, цифровые сигнальные процессоры, фильтрация сигнала, модель сигнала – гармоническая полуволна, составные полуволи, равномерные выборки, равномерная дискретизация времени, ступенчатая и линейная аппроксимация сигнала, допустимая погрешность.

1. Введение. Дискретизация времени и восстановление сигналов по выборкам является традиционной задачей обработки сигналов в

информационной технике, которая решается при проектировании или выборе каналов сбора и обработки информации в системах управления. Эта задача актуальна при проектировании аппаратуры обратной связи и каналов ввода данных для цифровых автоматических и автоматизированных систем управления реального времени (СРВ). Существенным свойством СРВ является тот факт, что они работают в «структуре времени, определяемой ходом внешних процессов» [1, 2] и в них применяют современные цифровые сигнальные процессоры (ЦСП) для решения задач первичной обработки сигналов [3–7].

Есть мнение, что полученные результаты позволяют упростить распределение задач и встраивание ЦСП в робастные СРВ при оптимизации их загрузки для первичной обработки сигналов.

1.1. Особенности первичной обработки сигнала цифровыми сигнальными процессорами. Физические параметры объекта управления снимаются датчиками в реальном времени, преобразуются в электрические сигналы, спектры которых характеризуются избыточными высокочастотными составляющими и эффектом алиасинга. Поэтому сигналы в СРВ подвергают предварительной обработке и ЦФНЧ - фильтрации при помощи цифровых сигнальных процессоров (ЦСП) в амплитудно-частотной области.

На рисунке 1 показан условно ЦСП и его основные функции при первичной обработке аналоговых сигналов в СРВ.

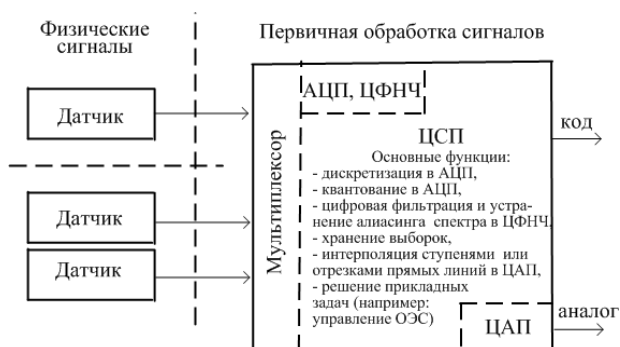


Рис. 1. Общая схема первичной обработки сигнала

Эти операции определяют требования к процедурам дискретизации времени и восстановления сигнала по выборкам [3–7], которые следуют из особенностей проектирования, применения и эксплуатации автономных СРВ с ЦСП:

– величина равномерного интервала дискретизации сигнала (функции) по времени находится по детерминированным выражениям, зависит от типа аппроксимации сигнала по выборкам и необходимой точности его восстановления. При этом адаптирующая аппроксимация с переменным шагом дискретизации и стохастические формулы, как правило, не применяются;

– аппроксимацию сигнала по выборкам из него осуществляют ступеньками или отрезками прямых линий и редко используют функцию $\text{sinc}(t)$ [5] по причине больших временных задержек;

– обработку информации с датчиков и выдачу сигналов управления на исполнительные органы и в контур отрицательной обратной связи СРВ выполняют максимум за один или два шага дискретизации времени на ЦСП;

– сохраняют значения сигнала без искажений в точках осуществления выборок из сигнала при его восстановлении.

Одним из основных физических параметров сигналов является частота среза f спектра [6, 7] или частота гармонической полуволны $f_c = 2f$, где индекс «с» означает признак «согласовано». Для предотвращения потери информации о сигнале при фильтрации сигнала в ЦФНЧ эту частоту принимают с превышением $f_c = 2Kf$. Коэффициент превышения $K \approx 8 \div 10$ определяют опытным путём. Для восстановления сигнала по равномерным выборкам в ЦСП СРВ применяют аппроксимацию ступенями или отрезками прямых линий.

К особым случаям можно отнести обработку сигналов (процесов), измерение частоты среза которых затруднено или отсутствует информация об их спектре. Это является неопределённостью объекта управления и его информационного описания. Такие средства и СРВ относят к робастным системам [3, 4] и применяют в разнообразных областях:

– геоакустическая эмиссия: одним из способов предсказания извержения вулканов является описание и систематизация геоимпульсов в амплитудно-временных координатах, выявление аномальных паттернов и их анализ для привязки к сейсмическим событиям [8];

– робототехника: для управления приводами манипулятора при копирующем типе управления [9], обеспечения перемещения схвата манипулятора из начального положения покоя в заданное конечное положение покоя за минимальное время [10];

– оптико-электронное наблюдение: для перехвата движущейся по предписанной траектории цели машиной Дубинса [11–14];

– нейронные сети: для обобщения модели химических взаимодействий между нейронами [15];

- ядерная медицина: для моделирования и дозиметрического планирования радиойодтерапии щитовидной железы [16];
- биотехнология, металлургия: для обеспечения качества продукта в условиях априорной неопределённости технологических параметров [6] процесса изготовления продукта и многое другое.

Модели процессов и сигналов, характеризующихся трудностями в определении спектра (СТОС), традиционно описывают в амплитудно-временной области: в автоматических системах график оптимального процесса управления задают в виде сопрягаемых парабол [17, 18]; моделируют равноускоренное движение манипулятора руки квадратичным трёхчленом [9, 10]; представляют сейсмические события в виде линейных степенных функций [8]. В качестве коэффициентов трёхчлена используют максимальные значения ускорения и скорости сигнала. Для сигналов СТОС определяют интервал дискретизации времени, используя два известных подхода: критерий наибольшего отклонения и, с существенным допущением, теорему НКШ. В настоящей статье предлагается новая компромиссная и точная модель.

1.2. Критерий наибольшего отклонения. Подход в виде критерия наибольшего отклонения (КНО) для определения величины интервала дискретизации времени T_p . [3, 6, 19] теоретически обосновал В.Н. Хлистунов, который применил для этой цели наибольшее значение остаточного члена формулы Лагранжа. КНО используют в измерительной технике и в других применениях ЦСП [3–7]. Достоинства применения критерия КНО для сигналов СТОС следующие:

Получены и обоснованы выражения КНО для определения интервалов дискретизации времени $T_{p1} = \Delta/V$ и $T_{p2} = (8\Delta/a)^{0.5}$, в которых индекс « p » означает параболическую (parabola) зависимость, $\max^{(0)}(t) = \Delta$ - максимально допустимая погрешность аппроксимации сигнала ступенькой или отрезком линии по двум соседним выборкам, $\max^{(1)}(t) = V$ - наибольшая скорость сигнала, $\max^{(2)}(t) = a$ - наибольшее ускорение сигнала.

Для восстановления сигнала $p(t)$ по выборкам из него через интервал времени $T_{p1} = \Delta/V$ используется ступенчатая аппроксимация. Для интервала дискретизации $T_{p2} = (8\Delta/a)^{0.5}$ используется линейная аппроксимация. Эти условия аппроксимации диктует формула Лагранжа.

Недостатки критерия наибольшего отклонения, которые не определены в [3, 6, 19]:

- Отсутствует связь величины граничной частоты f в спектре сигнала $p(t)$ с максимальными значениями параметров $\{V, a\}$ сигнала.
- Совместное рассмотрение выражений КНО T_{p1} и T_{p2} показывает, что T_{p1} зависит только от максимальной скорости V и погрешно-

сти Δ ступенчатой аппроксимации сигнала по выборкам, а T_{p2} зависит только от максимального ускорения (a) и погрешности Δ линейной аппроксимации сигнала по выборкам. Зависимость T_{p2} одновременно от скорости и ускорения сигнала В.Н. Хлистуновым не получена. Это обусловлено тем, что его подход к получению КНО, заключается в использовании остаточного члена формулы Лагранжа и не позволяет (принципиально) получить необходимый результат.

в) Отсутствует возможность определения величины T_{pn} (если $n \geq 3$) при условии восстановления сигнала при помощи ступенчатой или линейной аппроксимации, так как остаточный член формулы Лагранжа при $n = 3$ содержит максимальную величину третьей $\max p^{(3)}(t) = R$ производной от сигнала. При этом аппроксимация осуществляется квадратичными парабололами по трём точкам, а не отрезками прямых линий или ступеньками. Подобный недостаток препятствует определению величины шага T_{pn} дискретизации сигнала $p(t)$ по времени, если $n \geq 3$.

Недостаток б) КНО устранен автором в статье [20], где получено выражение (математическая модель) в виде обобщенного критерия наибольшего отклонения (ОКНО) для линейной аппроксимации сигнала $p(t)$ отрезком прямой линии по двум точкам

$$T_{p2} = \frac{2\Delta}{V} + \frac{V}{a},$$

где при $V = (2\Delta a)^{0.5}$ это выражение преобразуется в КНО $T_{p2} = (8\Delta/a)^{0.5}$, для ступенчатой аппроксимации берём $T_{p1} = 0.5T_{p2}$, $a \rightarrow \infty$, получаем КНО $T_{p1} = \Delta/V$. Выражение $V = (2\Delta a)^{0.5}$ названо условием согласования максимальных значений параметров $\{\Delta, V, a\}$ сигнала $p(t)$.

Общим недостатком применения степенных приближений и модели ОКНО является отсутствие математической связи с частотными характеристиками сигналов и процессов типа СТОС или с математическими моделями, использующими эти частотные характеристики.

1.3. Теорема Найквиста, Котельникова, Шеннона. Второй подход к определению интервалов дискретизации времени традиционно определен теоремой НКШ (Найквиста, Котельникова, Шеннона) [21-23]. Он основан на частотных свойствах сигналов и процессов, свободен от общего недостатка критериев КНО и ОКНО.

Теорема НКШ (теорема отсчетов) [21] доказана К. Шенноном на основе теоремы В.А. Котельникова [22, 6] и работы Г. Найквиста [23, 3] и используется в теории и технике связи:

Теорема НКШ: Если функция не содержит частот выше f_{Γ} , она полностью определяется своими мгновенными значениями в моменты, отстоящие друг от друга на $1/(2f)$ сек.

Интервал времени $1/(2f)$ назван К. Шенноном [21] интервалом Г. Найквиста [23]. Недостатки теоремы НКШ при решении задач в управляющих системах СРВ следуют из процедур её практического использования для дискретизации и восстановления сигналов [21]:

а) Восстановление сигнала происходит в течение большого времени T_{Φ} существования этого звукового сигнала с последовательным преобразованием $\sin c(t)$ и суммированием всех мгновенных значений (выборки), которые были сделаны из этой функции через интервалы времени кратные T_s . При этом заранее известно, что справедливо неравенство $T_{\Phi} \gg T_s$. В СРВ большие задержки при обработке сигналов во времени не допустимы.

Подробнее: в устройстве восстановления функции по ее отсчетам в каждой точке отсчета nT_s ($n=1, 2, 3, \dots$) на интервале времени T_{Φ} помещают импульс с амплитудой, равной величине каждого отсчета, на место этого отсчета, осуществляют последовательное преобразование этого импульса по передаточной характеристике фильтра нижних частот, результаты преобразований накапливают в течение интервала времени T_{Φ} и получают исходную функцию.

Математически процесс преобразования и восстановления сигнала $p(t)$ (звукового сигнала) описывается рядом В.А. Котельникова [22, 6], который подобен гармоническому ряду Фурье и является сложным с точки зрения количества необходимых вычислительных операций для его реализации. Выражения для ступенчатой и линейной аппроксимации (для восстановления сигнала по выборкам из него) существенно проще в математическом описании и реализации.

б) Применяемые в практике разработки и реализации СРВ методы ступенчатой и линейной аппроксимации не используются и не упоминаются при доказательстве теоремы НКШ [21, 22].

в) В теореме НКШ для определения T_s используется только граничная частота f в спектре частот сигнала, а остальные параметры $\max\{\Delta, V, a\}$ сигнала $p(t)$ не используются.

г) На практике [5] и в теории [3, 6] при использовании теоремы НКШ и ступенчатой или линейной аппроксимации сигнала по выборкам обычно используют эмпирическое соотношение $f_s = (8 \div 10)2f$, которое является приблизительным и не имеет строгого теоретического обоснования. Это обусловлено большой популярностью и известностью теоремы НКШ и делает погрешность расчётов неопределённой величиной.

В связи с этими и прочими недостатками нерационально определять интервал дискретизации сигнала T_s по теореме НКШ при осуществлении точных инженерных расчетов, при использовании ступенчатой и линейной аппроксимации для восстановления сигнала $p(t)$. В.Н. Хлистуновым [24] сделан справедливый вывод о необходимости использования теоремы НКШ только в теории связи (радиотехнике) для восстановления речевого сигнала, но не в системах измерения сигналов и управления в реальном времени.

1.4. Постановка задачи исследования. В настоящей статье предложена и рассмотрена модель сигнала в амплитудно-временной области в виде гармонической полуволны (и её новых модификаций). Описаны основные свойства модели, которые позволяют установить связь её параметров с амплитудно-частотными параметрами сигнала. Для этого получены выражения, связывающие математически два типа описания сигналов. Гармоническая полуволна является половиной волны частоты среза f спектра (повторяется с частотой $f_c = 2f$). Кроме наибольшего параметра f_c она характеризуется наибольшей скоростью сигнала $\max p^{(1)}(t) = V$, наибольшим ускорением сигнала $\max p^{(2)}(t) = a$. Амплитуда гармонической полуволны $\max p^{(0)}(t) = \Delta$ является максимально допустимой погрешностью аппроксимации сигнала ступенькой или отрезком линии по двум соседним выборкам. Далее покажем, что гармоническая полуволна является универсальной моделью для описания в амплитудно-частотной области и в амплитудно-временной области процессов и сигналов и связывает две группы максимальных величин параметров сигнала $\max\{\Delta, f_c\}$ и $\max\{\Delta, V, a, T_c\}$, где $T_c = 1/f_c$ – временной интервал полуволны. Первая версия подобной модели предложена и применена автором при разработке периферийных и сетевых устройств ЭВМ «Эльбрус-90микро» [25]. Развитием этой модели являются составные полуволны, которые являются общим случаем гармонических полуволн. При обработке сигналов, характеризующихся трудностями в определении спектра (СТОС) предлагается по замеренным (известным) параметрам такого сигнала $\max\{\Delta, V, a\}$ определить временной интервал полуволны $T_c = 1/f_c$. Для получения новых выражений (моделей) применим величины T_c или f_c и сформулированные ранее свойства и выражения для двух классов полуволн.

В качестве основного инструмента сравнения различных моделей полуволн используем условие согласования параметров сигнала (УСП). Максимальные параметры сигнала могут быть согласованными между собой или входить в УСП в виде равенства. Параметры сигнала могут входить в неравенство, тогда УСП будет отражать процент согласования. Покажем существование и исследуем взаимозависимость

коэффициента УСП и величины интервала дискретизации времени сигнала. Сравним величины интервалов дискретизации времени T_c согласованных гармонических полувольт [26] и предложенных новых моделей для согласуемых составных полувольт.

Модели и выражения будут полезны разработчикам СРВ для определения обоснованной величины интервала дискретизации времени сигнала (процесса) и, в связи с этим, для обеспечения оптимальной загрузки ЦСП при разработке СРВ с ЦСП.

2. Согласованная модель гармонической полувольты, её свойства и постановка задачи исследования

2.1. Два простейших частных случая нахождения шага дискретизации времени сигнала. Уравнение простейшей гармонической полувольты [26]:

$$p(t) = \Delta \sin \omega t, \quad (1)$$

где $0 \leq t \leq T_c$, $T_c = T/2 = 1/2f$ – интервал Найквиста, f – частота среза в спектре сигнала $p(t)$ после фильтрации в ЦФНЧ, T – период частоты среза, $\omega = 2\pi f$ – круговая частота, Δ – амплитуда гармонической полувольты или максимальная погрешность линейной аппроксимации двух выборок из полувольты в моменты времени $t = 0$ и $t = T_c$.

Первая производная ($i = 1$) от гармонической полувольты:

$$p^{(i=1)}(t) = \omega \Delta \cos \omega t, \quad (2)$$

где (i) – порядок производной, $1 \leq i \leq n$. Вторая ($i = 2$) и третья ($i = 3$) производные от гармонической полувольты:

$$p^{(i=2)}(t) = -\omega^2 \Delta \sin \omega t, \quad p^{(i=3)}(t) = -\omega^3 \Delta \cos \omega t. \quad (3)$$

Гармоническая полуволна и две производные от неё изображены на рисунке 2.

Выражение (1) (гармоническая полуволна) описывает математически половину $T_c = T/2 = 1/2f$ периода T частоты среза f в спектре сигнала $p(t)$. Частота $2f$ широко используется в теории связи, управления и в информатике [4–6].

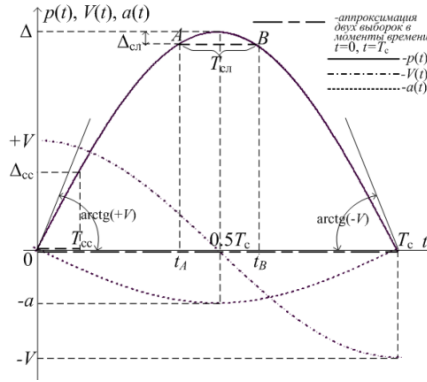


Рис. 2. Согласованная, гармоническая полуволна

Первая выборка из сигнала $p(t)$ осуществлена в момент времени $t = 0$, вторая – в момент времени $t = T_c = T/2$. Линейная интерполяция между двумя выборками изображена двух штриховой линией. Погрешность такой интерполяции Δ достигает максимального значения в точке $(0.5T_c, \Delta)$, которая является вершиной гармонической полуволны. Проходящая через вершину вертикальная линия $t = 0.5T_c = 0.25T$ служит её осью симметрии. Связь круговой частоты ω и интервала гармонической полуволны T_c запишем в виде выражения

$$\omega = 2\pi f = \frac{2\pi}{T} = \frac{\pi}{T_c}. \quad (4)$$

Определим интервал гармонической полуволны T_c через максимальные параметры сигнала, которые должны быть измерены при исследовании спектра сигнала $p(t)$ или получены при обработке результатов обследования объекта управления, если сигнал является СТОС.

В выражение первой производной (2) подставим моменты времени выборок из сигнала $p(t)$: $p(t = 0) = 0$ и $p(t = T_c) = 0$ и приравняем двум её максимальным значениям: $\max p^{(1)}(t = 0, t = T_c) = |\pm V|$, учтём (4), получим:

$$T_{c1} = \pi \frac{\Delta}{V}, \quad (5)$$

где индекс $(i = 1)$ из (2) переходит в индекс T_{c1} при получении (5).

В выражение второй производной (3) подставим момент времени $(t = 0.5T_c)$, при котором наступает максимум ускорения сигнала

$p(t)$ и приравняем её максимальному значению: $\max p^{(2)}(t = 0.5T_c) = |-a|$. Используем (4), получим цепочку равенств:

$$T_{c1} = \pi \frac{\Delta}{V} = T_{c2} = \pi \frac{V}{a} = T_{c2} = \pi \sqrt{\frac{\Delta}{a}}. \quad (6)$$

Аналогично (6), при измеренном значении $\max p^{(3)}(t = 0, t = T_c) = |\pm R|$, можно записать более длинную цепочку равенств:

$$T_{c1} = \pi \frac{\Delta}{V} = T_{c2} = \pi \frac{V}{a} = T_{c2} = \pi \sqrt{\frac{\Delta}{a}} = T_{c3} = \pi \frac{a}{R} = T_{c3} = \pi \sqrt[3]{\frac{\Delta}{R}}. \quad (7)$$

2.2. Общие выражения для нахождения интервала гармонической полуволны. Применим математическую индукцию к выражениям (5), (6), (7) и запишем цепочку равенств в общем виде:

$$T_{ci} = \pi \frac{\max p^{(i-1)}(t)}{\max p^{(i)}(t)} = \pi \left[\frac{\Delta}{\max p^{(i)}(t)} \right]^{1/i}, \quad (8)$$

где $1 \leq i \leq n$, $\max p^{(0)}(0.5T_c) = \Delta$, $\max p^{(1)}(0) = V$, $\max p^{(2)}(0.5T_c) = -a, \dots$ Преобразование первых пар звеньев цепочек равенств (6) и (7) даёт условие согласования максимальных параметров (УСП) сигнала: $V^2 = \Delta a$, или условие справедливости равенств звеньев цепочек между собой. Можно записать общее выражение для определения УСП:

$$\max p^{(i)}(t) = \sqrt{\max p^{(i-1)}(t) \times \max p^{(i+1)}(t)}, \quad (9)$$

где $1 \leq i \leq (n-1)$, $n \geq 2$. Из второй части выражения (8) и выражения (4) запишем выражение для определения максимальных временных параметров гармонической полуволны:

$$\max p^{(i)}(t) = \Delta (2\pi f)^i \text{ или } \omega = 2\pi f = \left(\frac{\max p^{(i)}(t)}{\Delta} \right)^{1/i}, \quad (10)$$

где при $(i = 1) \omega = V/\Delta$, при $(i = 2) \omega = (a/\Delta)^{1/2}$, при $(i = 3) \omega = (R/\Delta)^{1/3}$.

2.3. Свойства согласованных гармонических полувольт. Записанные выражения и рисунок 2 позволяют сформулировать основные свойства моделей, которые используем далее:

Свойство 1. (Аксиома математического подобия) Гармоническая полуволна (1) сигнала $p(t)$ имеет полный комплекс частотно-временных и амплитудных параметров $\{f_c, \Delta, \max p^{(i)}(t), \text{ где } 0 \leq i \leq n\}$, которые имеет сигнал $p(t)$, поэтому может быть использована как математическая модель для исследования зависимостей максимальных величин параметров сигнала $p(t)$ между собой.

Свойство 2. (Аксиома линейной аппроксимации). Любой сигнал $p(t)$, состоящий из гармоник частот от 0 до f , можно восстановить при помощи линейной аппроксимации с максимальной погрешностью Δ равной амплитуде гармоники частоты f при помощи выборок, следующих друг за другом через $1/(2f)$ секунд, это следует из описания модели гармонической полуволны (1) и рисунка 2.

Свойство 3. В области максимального ускорения гармонической полуволны $[\max p^{(2)}(t = 0.5T_c) = a]$ относительная погрешность линейной аппроксимации соседних выборок из этой полуволны максимальна, относительная погрешность ступенчатой аппроксимации соседних выборок из этой полуволны минимальна (1), (2), (3), [6].

Свойство 4. В областях максимальной скорости гармонической полуволны $[\max p^{(1)}(t=0, t=T_c)=\pm V]$ относительная погрешность линейной аппроксимации соседних выборок из этой полуволны минимальна, относительная погрешность ступенчатой аппроксимации соседних выборок из этой полуволны максимальна (1), (2), (3), [6].

Свойство 5. Условие согласования параметров (9) является равенством, это указывает на факт: любая измерительная погрешность в замерах параметров амплитуды Δ и частоты f , приводит к искажению величины интервала гармонической полуволны (8).

Свойство 6. Ступенчатая и линейная аппроксимации соседних выборок в области максимального ускорения $[\max p^{(2)}(t = 0.5T_c) = a]$ имеют равные значения погрешности аппроксимации, если частота повторения выборки, аппроксимируемой ступенью в два раза выше частоты выборки, аппроксимируемой отрезком прямой линии.

Свойство 7. Максимальная производная порядка (i) от сигнала $p(t)$ или гармонической полуволны является средним геометрическим максимальных производных порядков $(i - 1)$ и $(i + 1)$ (9).

Свойство 8. Если измерить параметры гармоники частоты среза спектра сигнала: амплитуду Δ и частоту f и подставить их значения в первый вариант (10), то можно рассчитать $(1 \leq i \leq n)$ величин параметров $\max p^{(i)}(t)$ сигнала $p(t)$ или гармонической полуволны.

Свойство 9. (обратное Свойству 8). Если амплитуду Δ и частоту f гармонической полуволны замерить затруднительно из-за отсутствия информации о спектре (СТОС), то можно измерить максимально допустимое отклонение Δ объекта от двух соседних выборок, максимальную скорость V отклонения и по второму варианту (10) определить эквивалентную граничную круговую частоту ω .

2.4. Основные недостатки модели гармонической полуволны, направления их устранения или задачи исследования.

Недостаток 1. Например, поверхность УСП $V = (\Delta a)^{0.5}$, изображённая на рисунке 3 иллюстрирует Свойство 5.

Для устойчивого согласования параметров сигнала необходимо, чтобы эта поверхность делила пространство Δ, V, a на две части по признаку согласования и рассогласования параметров сигнала. Тогда разрешенная комбинация в виде $0 \leq V \leq (\Delta a)^{0.5}$ будет находиться в полупространстве ниже поверхности $V = (\Delta a)^{0.5}$.

Покажем далее, что такой результат даёт исследование зависимости величин интервалов двух вариантов составных полуволн от величины коэффициента согласования параметров (УСП) этих моделей.

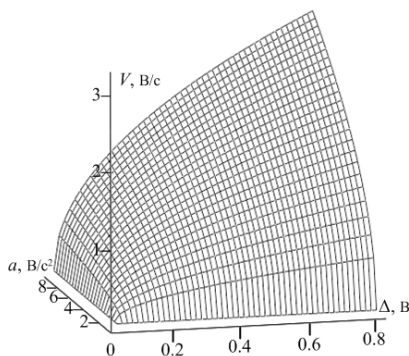


Рис. 3. Место точек согласованных параметров сигнала (процесса) в виде параболической поверхности

Недостаток 2. Из аксиомы по Свойству 2 и выражения (10) следует: величина максимальной погрешности Δ (после восстановления сигнала $p(t)$ по выборкам при помощи ступенчатой и линейной аппроксимации) равна амплитуде Δ гармоники частоты среза f в спектре частот сигнала $p(t)$, которая осталась в составе сигнала после низкочастотной фильтрации и может быть достаточно большой.

Этот недостаток не позволяет обосновать или уточнить эмпирическое соотношение $T_c \leq (1/2f)/(8 \div 10)$ для линейной интерполяции, которое явно приводит к уменьшению интервала дискретизации по сравнению с интервалом гармонической полуволны $(1/2f)$ в $(8 \div 10)$ раз.

Используем Свойство 1 и процедуру равномерной дискретизации времени интервалов гармонической полуволны и интервалов предлагаемых, составных полуволн для получения математических выражений, заменяющих и уточняющих это соотношение.

3. Представление моделей составных полуволн, согласуемых по величине максимальной скорости сигнала.

3.1. Модель составной полуволны при постоянной скорости на начальном и конечном её участках. Для устранения Недостатка 1 используем составную полуволну, графическая модель которой изображена на рисунке 4. Предположим, что для сигнала или процесса $p(t)$ при обследовании объекта управления измерены три параметра Δ , V_C , a , при этом величина максимальной скорости $V_C = \text{const} \leq V = (\Delta a)^{0.5}$, а частота f неизвестна из-за свойства сигнала СТОС. Другими словами УСП сигнала не выполняется, а параметры гармонической полуволны являются рассогласованными. Для согласования траектории используем коэффициент k и согласуем наибольшую скорость изменения V_C с параметрами Δ , a сигнала:

$$V_C = kV = k(\Delta a)^{0.5}, \quad (11)$$

где $0 \leq k \leq 1$. Из выражения (11) $V_C = \text{const} \leq V = (\Delta a)^{0.5}$ следует:

Свойство 10. *Трёхмерная область $0 \leq V(t) \leq (\Delta a)^{0.5}$ располагается ниже поверхности $V = (\Delta a)^{0.5}$ (см. рисунок 3), включает в себя эту поверхность и является областью существования модели траектории составной полуволны.*

Графическая модель траектории составной полуволны характеризуется максимальной скоростью V_C , которая постоянна в начале и конце траектории $\pm V_C = \pm \text{const}$. Поэтому эту полуволну назовём «полуволной, согласуемой двумя отрезками прямых линий» и введём индекс «л» в обозначение величины интервала составной полуволны $T_{л}$.

Первая выборка из составной полуволны осуществлена в момент времени $(-t_{к0})$, вторая – в момент времени $(T_c + t_{к0})$ (см. рисунок 4).

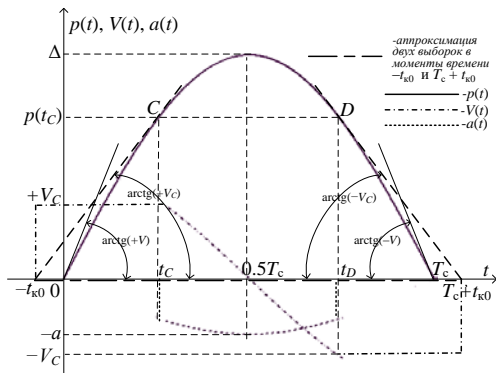


Рис. 4. Составная полуволна, согласуемая двумя отрезками прямых линий, $\pm V_C = \pm \text{const}$

Таким образом, интервал составной полуволны:

$$T_{\text{л}} = T_c + 2|t_{\text{к0}}|. \quad (12)$$

Между двумя выборками сигнал отклоняется по составной полуволне, состоящей из трёх частей:

- на интервале времени $(-t_{\text{к0}} \leq t \leq t_c)$ с постоянной скоростью $V_C \leq (\Delta a)^{0.5}$, по прямой линии касательной к гармонической полуволне синусоиды (1) в точке C;
- между точками касания C и D линий касательных к синусоиде (1) по гармонической полуволне (1);
- на интервале времени $(t_D \leq t \leq T_c + t_{\text{к0}})$ с постоянной скоростью $-V_C \geq -(\Delta a)^{0.5}$, по прямой линии касательной к (1) в точке D.

Для нахождения интервала полуволны $T_{\text{л}}$ по выражению (12) определим момент времени $t_{\text{к0}}$ и найдём касательную к полуволне синусоиды (1) в точке C. Первую производную от (1) в точке C приравняем скорости V_C , учтём (11) $V_C = kV$, получим:

$$t_c = \frac{1}{\omega} \arccos k \quad \text{и} \quad p(t_c) = \Delta \sqrt{1 - k^2}.$$

Пересечение касательной с осью координат $(0, t)$ на рисунке 4 даёт:

$$t_{\text{к0}} = \frac{1}{\omega} \left[\sqrt{\frac{1}{k^2} - 1} - \arccos k \right].$$

Преобразуем это выражение – используем (4) $\omega = \pi/T_c$ и (11) $V_C = kV$. Подставим преобразованное $t_{к0}$ в (12) и получим:

$$T_{л} = T_c \left[1 + \frac{2}{\pi} \left(\sqrt{\frac{1}{k^2} - 1} - \arccos k \right) \right], \quad (13)$$

где $0 \leq k \leq 1$. Для сравнения величин интервалов для согласованной гармонической полуволны T_c и согласуемой линией составной полуволны $T_{л}$ (13) введем относительные величины интервалов рассматриваемых полуволн $\tau_{л} = T_{л} / T_c$ и $\tau_c = T_c / T_c = 1$. Преобразуем (13), получим графики $\tau_{л}$ и τ_c (см. рисунок 5).

Графики показывают, что при уменьшении коэффициента согласования k или уменьшении максимальной скорости изменения сигнала $V_C = kV$ происходит существенное увеличение интервала времени составной полуволны $T_{л}$. Этот результат позволяет сформулировать новое информационное свойство:

Свойство 11. Уменьшение величины коэффициента k или увеличение степени рассогласования величин временных максимальных значений параметров $\max\{\Delta, V, a\}$ сигнала или процесса $p(t)$ приводит к увеличению интервала времени $T_{л}$ составной полуволны. При этом максимальная абсолютная погрешность измерения равна амплитуде гармонической полуволны Δ или неизменна.

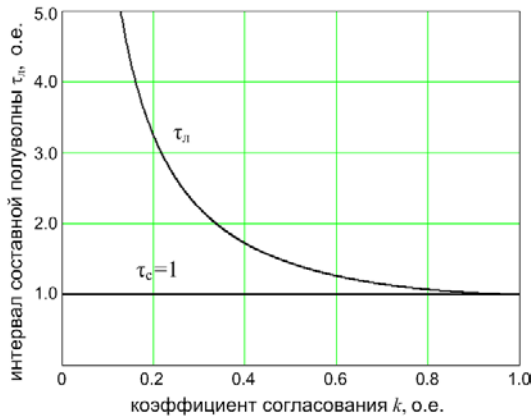


Рис. 5. Зависимость относительного интервала составной полуволны от коэффициента k согласования параметров сигнала

Для проверки общности моделей (13) и (6) зададим $k = 1$ в (13) или выполним условие согласования. Получим $T_{\Delta}(k = 1) = T_C$, что позволяет сформулировать свойство:

Свойство 12. Математическая модель гармонической полуволны с согласованными временными параметрами сигналов (процессов) является частным случаем модели составной полуволны, согласуемой двумя симметричными отрезками касательных прямых.

3.2 Модель составной полуволны при переменной скорости на начальном и конечном участках. Графическая модель составной полуволны с переменной скоростью изображена на рисунке 6.

Согласующие участки этой полуволны описываются двумя симметричными частями гармонической функции, где нижний индекс «к» обозначает касание левой части гармонической функции в точке С к гармонической полуволне (1):

$$p(t) = \Delta_k \sin(\omega_k t + \varphi_0), \tag{14}$$

где $\Delta_k, \varphi_0, \omega_k$ – соответственно, амплитуда, фаза и круговая частота составной полуволны, $(-t_{k0}) \leq t \leq 0.5T_C$, $\varphi_0/\omega_k = t_{k0}$, $\omega_k = 2\pi/T_k$, $\omega_k = V_k/\Delta_k$.

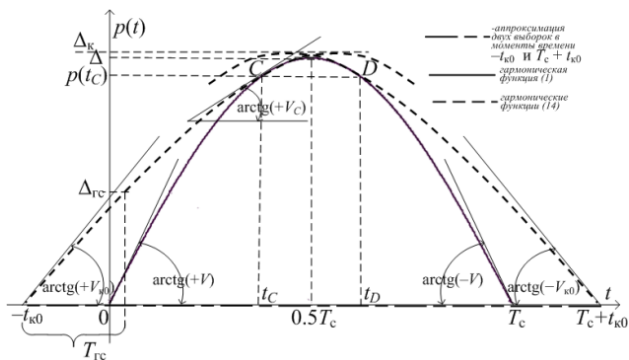


Рис. 6. Составная полуволна, согласуемая двумя гармоническими линиями, $\pm V(t) = \pm var$

Такая полуволна согласована гармоникой, поэтому введём индекс «г» в обозначение интервала времени полуволны T_{Γ} .

Предположим, что для сигнала или процесса $p(t)$ получены три параметра Δ, V_C, a . Причём величина максимальной скорости в точке С равна $V_C = kV = k(\Delta a)^{0.5}$, где $0 \leq k \leq 1$. Процесс отклонения полувол-

ны от прямой, аппроксимирующей две соседние выборки в моменты времени $(-t_{к0})$ и $(T_c + t_{к0})$, происходит по трём участкам:

- в интервале времени $(-t_{к0} \leq t \leq t_c)$, по левой части гармоники (14) от точки $(-t_{к0})$ до точки C , при этом скорость $V_C \leq V_{к0} \leq V = (\Delta a)^{0.5}$;
- между точкой касания C гармонической функции (14) к синусоиде (1) в момент времени $t = t_c$ до точки её пересечения с осью симметрии в момент времени $t = 0.5T_c$;
- в интервале времени $(0.5T_c \leq t \leq T_c + t_{к0})$ симметрично второму и первому участкам относительно оси симметрии $t = 0.5T_c$.

Интервал времени составной полуволны T_Γ можно определить по выражению (12) с учётом величины фазы (14) гармонической траектории наибольшего отклонения $|t_{к0}| = \varphi_0/\omega_k$ и $\omega_k = 2\pi/T_k$:

$$T_\Gamma = T_c + \frac{1}{\pi} T_k \varphi_0 = T_c \left(1 + \frac{2}{\pi} \varphi_0 r \right), \quad (15)$$

где $r = T_k/2T_c$, $T_k \geq T = 2T_c$, $r \geq 1$. Для определения φ_0 воспользуемся равенством скоростей V_C в точке взаимного касания $C(t_c, p(t_c))$ гармонических функций (1) и (14). От этих выражений возьмём производные по времени, используем выражения подобные (11) $V_C = kV$, $V_C = mV_{к0}$. Применим выражение (4) $\omega = \pi/T_c$ и ему подобное выражение $\omega_k = 2\pi/T_k$, получим:

$$\varphi_0 = \arccos m - \frac{1}{r} \arccos k. \quad (16)$$

Найдём связь между величинами коэффициентов согласования параметров сигнала k , m и коэффициентом r .

Выражение (1) при $t = t_c$ запишем в виде $\Delta \sin \omega t_c = p(t_c)$ и последовательно преобразуем его по известной тригонометрической формуле $\arcsin x = \arccos[(1 - x^2)^{0.5}]$, получим:

$$\omega t_c = \arcsin \left[p(t_c)/\Delta \right] = \arccos \left[1 - (p(t_c)/\Delta)^2 \right]^{0.5}. \quad (17)$$

Первая производная от (1) при $t = t_c$ равна скорости $V(t_c) = V_C$, из (2) запишем: $\Delta \omega \cos \omega t_c = V_C$. В это выражение подставим вторую часть выражения (17) и применим тригонометрическую формулу $\cos \arccos x = x$. Получим выражение $p^2(t_c) = \Delta^2 - (V_C/\omega)^2$, в которое последовательно подставим $\omega = V/\Delta$ из (10) и $V_C = kV$ из (11), в итоге получим: $p^2(t_c) = \Delta^2(1 - k^2)$.

Аналогично преобразуем выражение (14), при этом используем $\omega_k = V_{k0}/\Delta_k$ из (10), и $V_C = mV_{k0}$ из (11). Получим выражение: $p^2(t_C) = \Delta_k^2(1 - m^2)$. Два выражения для нахождения $p^2(t_C)$ преобразуем совместно и получим:

$$\frac{\Delta^2}{\Delta_k^2} = \frac{1 - m^2}{1 - k^2} = M^2. \quad (18)$$

Определим отношение амплитуд составной траектории отклонения Δ и Δ_k через коэффициенты согласования параметров сигнала m , k и коэффициент r . Для этого последовательно применим выражения (10) $\Delta = V/\omega$, $\Delta_k = V_{k0}/\omega_k$, выражения (4) $\omega = \pi/T_C$, $\omega_k = 2\pi/T_k$, выражения $V_C = kV$, $V_C = mV_{k0}$, получим частное $\Delta/\Delta_k = m/kr$. Подставим это выражение в (18) и выделим коэффициент согласования r , который в соответствии с (15) $r \geq 1$:

$$r = \frac{m}{k} \sqrt{\frac{1 - k^2}{1 - m^2}} = \frac{m}{kM} \geq 1. \quad (19)$$

Из (19) следует: $m \geq k$. Выражение (19) умножим на (16) ($\varphi_0 \times r$) и результат подставим в (15). Введем относительную величину интервала гармонической составной полуволны $\tau_r = T_r/T_C$. В результате преобразований получим окончательное выражение:

$$\tau_r = \frac{T_r}{T_C} = 1 + \frac{2}{\pi} \left(\frac{m}{k} \sqrt{\frac{1 - k^2}{1 - m^2}} \arccos m - \arccos k \right), \quad (20)$$

где $0 \leq m \leq 1$, $m \geq k$. Построим графики τ_r для $m = 0.2$, $m = 0.5$, $m = 0.8$ и для наглядности совместим их с графиками τ_l и τ_c (см. рисунок 7).

Каждый график τ_r определён в своей верхней части ($\tau_r \geq \tau_c = 1$), при коэффициенте согласования из замкнутого интервала $0 \leq k \leq m$, что следует из (19) и (15). Правые границы определения лежат в точках пересечения графика $\tau_c = 1$ с каждым из трёх графиков зависимостей $\tau_r = F(k, m = \text{const})$.

Найдём предел выражения (20) $\lim_{\tau_r(m \rightarrow 1)} = 1$, по правилу Лопиталя, при этом раскроем неопределённость $0/0$, получим выражение (13). Из этого следует свойство:

Свойство 13. Математическая модель составной полуволны с согласованными временными параметрами сигналов (процессов) при

помощи двух отрезков прямых линий является частным случаем модели составной полуволны, согласуемой двумя симметричными частями гармонической функции.

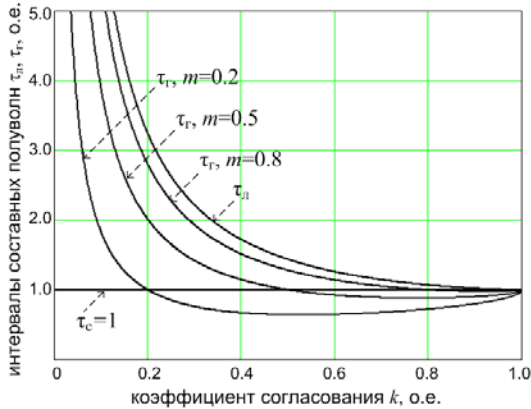


Рис. 7. Зависимости относительных интервалов составных полуволн от коэффициента k согласования параметров сигнала

4. Модели дискретизации полуволн по времени. Для избыточной фильтрации сигнала применяют эмпирическое соотношение $T_c \leq (1/2f)/(8 \div 10)$, которое явно приводит к уменьшению интервала дискретизации по сравнению с интервалом Найквиста (см. Недостаток 2, пункт 2.4). Кроме этого по Свойствам 3, 4 относительная максимальная погрешность аппроксимации соседних выборок из описанных полуволн зависит от типа аппроксимации. Учтём в дальнейшем эти обстоятельства при описании моделей.

4.1. Модель дискретизации гармонической полуволны и восстановление её ступеньками. Гармоническая полуволна согласована, имеет УСП ($V = (\Delta a)^{0.5}$), изображена на рисунке 2. Первая выборка из полуволны произведена в момент времени $t = 0$. Предположим, что вторая выборка произведена не в $t = T_c$, а ранее: в момент времени $t = T_{cc}$, где первый индекс «с» означает согласованность параметров полуволны, второй индекс «с» означает аппроксимацию сигнала по выборкам ступеньками. При этом абсолютная погрешность ступенчатой аппроксимации сигнала или процесса $p(t)$ составляет Δ_{cc} . Используем Свойство 4. Связь времени и максимальной погрешности получим подстановкой в выражение (1) указанных величин: $\Delta \sin \omega T_{cc} = \Delta_{cc}$. Обозначим $\delta_{cc} = \Delta_{cc}/\Delta$, используем (4) $\omega = \pi/T_c$, получим

интервал дискретизации полуволны $T_{cc} = (T_c/\pi)\arcsin\delta_{cc}$. Погрешность аппроксимации δ_{cc} выразим через коэффициент $Q_{cc} = T_c/T_{cc}$ превышения частоты гармонической полуволны частотой равномерной дискретизации этой полуволны или сигнала при восстановлении сигнала ступеньками:

$$\delta_{cc} = \sin \frac{\pi}{Q_{cc}}. \tag{21}$$

На рисунке 8 изображён график δ_{cc} для модели по выражению (21).

4.2. Модель дискретизации составной полуволны, согласованной отрезками прямых линий. Траектория изображена на рисунке 4. При обследовании объекта управления измерены три параметра (Δ, V_C, a) , при этом $(V_C \leq (\Delta a)^{0.5})$ или УСП сигнала не выполняется. Первая выборка из полуволны произведена в момент времени $t = -t_{к0}$, вторая выборка произведена не в момент времени $t = (T_c + t_{к0})$, а ранее: через интервал времени $T_{лс}$ от первой выборки.

Из рисунка 4 следует $T_{лс} = \Delta_{лс}/V_C$. Зададимся необходимой относительной погрешностью аппроксимации $\delta_{лс}$: $\Delta_{лс} = \delta_{лс}\Delta$. Используем Свойство 4. Используем (11) $V_C = kV$, (10) $\Delta/V = \omega$, (4) $\omega = \pi/T_c$. В итоге получим: $T_{лс} = \delta_{лс}/k\omega$.

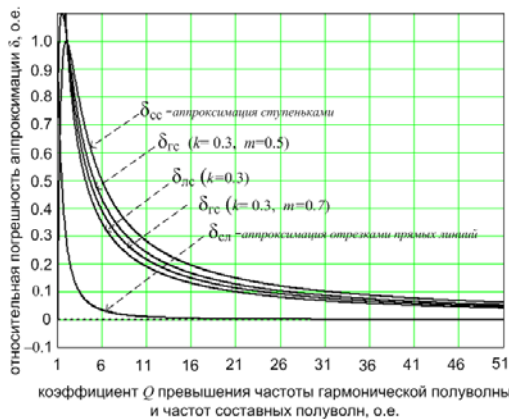


Рис. 8. Зависимость погрешности δ ступенчатой и линейной аппроксимации от коэффициента Q превышения частоты гармонической полуволны и частот составных полуволн частотой дискретизации сигнала

Для записи коэффициента $Q_{\text{лс}} = T_{\text{л}}/T_{\text{лс}}$ превышения частоты гармонической полуволны частотой равномерной дискретизации сигнала при восстановлении сигнала ступеньками воспользуемся (13), (4) $\omega = \pi/T_{\text{с}}$. Погрешность аппроксимации $\delta_{\text{лс}}$ выразим через коэффициент $Q_{\text{лс}}$:

$$\delta_{\text{лс}} = \frac{k}{Q_{\text{лс}}} \left(\pi + 2 \sqrt{\frac{1}{k^2} - 1} - 2 \arccos k \right). \quad (22)$$

На рисунке 8 изображён график $\delta_{\text{лс}}$ при $k = 0.3$ для модели по выражению (22).

4.3. Модель дискретизации составной полуволны, согласованной частями гармонической функции. Траектория изображена на рисунке 6. При обследовании объекта управления измерены три параметра (Δ , $V_{\text{с}}$, a), при этом ($V_{\text{с}} \leq (\Delta a)^{0.5}$) или УСП сигнала не выполняется. Первая выборка из полуволны произведена в момент времени $t = -t_{\text{к0}}$, вторая выборка из полуволны произведена не в момент времени $t = (T_{\text{с}} + t_{\text{к0}})$, а через интервал времени $T_{\text{гс}}$ от первой выборки. Из рисунка 6: $\delta_{\text{гс}} = \Delta_{\text{гс}}/\Delta$, из (18) $M = \Delta/\Delta_{\text{к}}$. В выражении (14) примем $\varphi_0 = 0$ или сдвинем начало гармонической функции в начало координат. Это упростит получаемые выражения и не повлияет на величину погрешности аппроксимации ступенькой. Определим из (14): $\sin(\omega_{\text{к}} T_{\text{гс}}) = \delta_{\text{гс}} M$, учтём из (14) и (15): $\omega_{\text{к}} = 2\pi/T_{\text{к}} = \pi/rT_{\text{с}}$. Запишем $T_{\text{гс}} = (rT_{\text{с}}/\pi) \arcsin(\delta_{\text{гс}} M)$.

Введём коэффициент $Q_{\text{гс}} = T_{\text{г}}/T_{\text{гс}}$ превышения частоты составной полуволны, согласованной двумя гармоническими функциями, частотой её равномерной дискретизации, используем (20), получим:

$$\delta_{\text{гс}} = \sqrt{\frac{1-k^2}{1-m^2}} \sin \left(\frac{k}{mQ_{\text{гс}}} \sqrt{\frac{1-m^2}{1-k^2}} \left(\pi + 2 \frac{m}{k} \sqrt{\frac{1-k^2}{1-m^2}} \arccos m - 2 \arccos k \right) \right), \quad (23)$$

где $0 \leq k \leq 1$, $0 \leq m \leq 1$, $m \geq k$. На рисунке 8 изображены два графика $\delta_{\text{гс}}$ при $k = 0.3$, $m = 0.7$ и при $k = 0.3$, $m = 0.5$ для модели (23).

4.4. Модель дискретизации гармонической полуволны при восстановлении сигнала отрезками прямых линий. Воспользуемся рисунком 2 и рассмотрим отрезок прямой $[A, B]$, который аппроксимирует полуволну Котельникова по двум соседним выборкам из нее в моменты времени t_A и t_B . Из рисунков 2, 4 и 6 следует:

Свойство 14. Все рассмотренные в этой статье варианты полувольты с согласованными и согласуемыми параметрами имеют одинаковую относительную погрешность линейной аппроксимации, так как для них справедливы Свойства 3 и 10 или все они имеют одинаковую геометрию вершины полувольты.

В соответствии со Свойством 14 на рисунке 8 приведён только один график с погрешностью от линейной аппроксимации $\delta_{\text{сл}} = \Delta_{\text{сл}}/\Delta$, где индекс «с» означает согласованность параметров гармонической полувольты, второй индекс «л» означает аппроксимацию гармонической полувольты или сигнала по выборкам отрезками прямых линий.

Отрезок $[A, B]$ является горизонтальным участком аппроксимации и делится осью симметрии $t = 0.5T_c$ на две равные части. Определим момент времени t_A начала отрезка. Воспользуемся коэффициентом $Q_{\text{сл}} = T_c/T_{\text{сл}}$ превышения частоты гармонической полувольты:

$$t_A = 0,5T_c - 0,5T_{\text{сл}} = 0,5T_c(1 - 1/Q_{\text{сл}}). \quad (24)$$

В выражение (1) подставим t_A (24), учтем $\omega = \pi/T_c$ из (4), получим $p(t_A) = \Delta \sin[0.5\pi(1 - 1/Q_{\text{сл}})] = \Delta \cos(\pi/2Q_{\text{сл}})$ – выражение горизонтального отрезка прямой, проходящей через точки A, B (см. рисунок 2). Зависимость погрешности аппроксимации $\Delta_{\text{сл}}(t)$ от времени t определим как разность выражения (1) и $p(t_A)$:

$$\Delta_{\text{сл}}(t) = \Delta \sin \omega t - \Delta \cos(\pi / 2Q_{\text{сл}}). \quad (25)$$

Выражение для максимального значения относительной погрешности восстановления $\delta_{\text{сл}} = \Delta_{\text{сл}}/\Delta$ получим подстановкой $t = 0.5T_c$ и (4) $\omega = \pi/T_c$ в (25):

$$\delta_{\text{сл}} = 1 - \cos(\pi / 2Q_{\text{сл}}). \quad (26)$$

На рисунке 8 изображен график $\delta_{\text{сл}}$ для модели по выражению (26), которая соответствует аппроксимации отрезками прямой линии равномерных выборок из сигнала или гармонической полувольты.

5. Обработка результатов моделирования. Для сравнения вариантов рассмотренных моделей приведена таблица 1 значений зависимости погрешности $\delta_{\text{сс}}$ и $\delta_{\text{сл}}$ аппроксимации от коэффициентов $Q_{\text{сс}}$ и $Q_{\text{сл}}$ для гармонической полувольты и также для составных полувольт.

Таблица 1. Связь коэффициента превышения частоты четырёх типов полувольт с величиной погрешности аппроксимации

		$Q_{cc} = Q_{lc} = Q_{rc} = Q_{cl}$, о.е.						
		6	11	16	21	26	31	36
δ , о.е.	«сс»	0.5000	0.2620	0.1960	0.1210	0.1210	0.1010	0.0870
	«лс»	0.3480	0.1900	0.1310	0.1000	0.0800	0.0670	0.0580
	«рс»	0.4330	0.2410	0.1660	0.1270	0.1030	0.0860	0.0740
	«сл»	0.0340	0.0100	0.0048	0.0028	0.0018	0.0013	0.0010

Из таблицы 1 (см. строку 4) следует, что эмпирическое соотношение $T_{cl} \leq (1/2f)/(8 \div 10)$ для линейной аппроксимации дает погрешность восстановления более 1%. Выражение $T_{cl} \leq (1/2f)/(Q_{cl}=11)$ соответствует погрешности в 1%. Для обеспечения погрешности в 0.1% необходимо использовать выражение $T_{cl} \leq (1/2f)/(Q_{cl}=36)$.

Из выражения (21) и таблицы 1 (см. строку 1) следует, что эмпирическое соотношение $T_{cc} \leq (1/2f)/(8 \div 10)$ для аппроксимации ступеньками дает погрешность восстановления более 30.9 %.

При использовании ступенчатой аппроксимации обнаруживается существенное превышение Q_{cc} частоты гармонической полувольты для всех типов наибольших отклонений над Q_{cl} (см. таблицу 2):

Таблица 2. Превышение частоты дискретизации четырёх типов полувольт для трёх типовых значений погрешности аппроксимации

$\min \delta$, %	Q_{cc} , о.е.	Q_{lc} , о.е.	Q_{rc} , $m=0.7$	Q_{cl} , о.е.
10.0	31.363	20.907	23.839	3.483
1.00	314.154	209.069	238.610	11.095
0.10	3142.000	2091.000	2386.000	35.121

Результаты, приведённые в таблицах 1 и 2, определяют:

Свойство 14. Для выполнения точных преобразований сигналов с датчиков СВВ с ЦСС при минимальном количестве выборок выгодно применять только линейную аппроксимацию для восстановления сигналов, для чего расчёт величин интервалов равномерной дискретизации времени нужно проводить по выражениям (26) и (6): $T_{cl} = T_{c2}/Q_{cl}$.

6. Пример использования моделей. Рассмотрим две авиационные оптико-электронные системы (ОЭС) производителей России и Англии [27] с заведомо различными по величине паспортными параметрами для проверки достоверности моделей. Найдём частоту опроса датчика азимута ОЭС для обеспечения заданной погрешности аппроксимации равномерных выборок. По полученным выражениям оценим возможность применения моделей по назначению и сравним их между собой. Особенностью ОЭС является отсутствие спектра сигнала ази-

мута, поэтому используем максимальные паспортные параметры: скорость V и ускорение вращения a (см. столбцы 2 и 3 таблицы 3) [27]. Согласованная модель гармонической полуволны возникает в режиме реверса ОЭС, который выполняется до реализации алгоритмов машины Дубинса [11]. Основными этапами реверса являются:

- ОЭС на максимальной угловой скорости ($+V$, рад/с) следит за воздушным объектом ВО1;
- ОЭС переключают на режим слежения за ВО2, направление вектора скорости которого противоположно вектору движения ВО1;
- включается максимальное отрицательное угловое ускорение ($-a$, рад/с²) и через интервал времени (6) $0.5T_c = 0.5\pi V/a$ (с) мгновенная скорость будет равна нулю $V(0.5T_c) = 0$ (см. рисунок 1), угол «мгновенная» останова ОЭС (из (9) при $i = 1$) будет равен $\Delta = V^2/a$ (см. столбцы 6 и 7 таблицы 3);
- далее скорость вращения возрастает до величины $V(T_c) = -V$. Процесс подобен гармонической полуволне (см. рисунок 2, Свойство 9), поэтому полуволна (1) применена для моделирования ОЭС.

Таблица 3. Паспортные параметры, исходные данные и результаты моделирования для двух типов ОЭС

Марка ОЭС	Максимальные, паспортные параметры ОЭС [27]		Пример исходных данных моделей (заданы)		Расчётные характеристики ОЭС (6), (9) при $i = 1$		Результаты расчётов параметров гармонической полуволны и параметров составной полуволны с восстановлением сигнала по выборкам ступенями и отрезками прямых линий					
	Угловая скорость V , рад/с	Угловое ускорение a , рад/с ²	δ , о.е.	k , о.е.	Интервал реверса $T_{c2} = \pi V/a$, с	Угол останова $\Delta = V^2/a$, рад	сс (21)		лс (22)		сл (26)	
							$Q_{лс}$, о.е.	$T_{лс}$, ме	$Q_{лс}$, о.е.	$T_{лс}$, ме	$Q_{сл}$, о.е.	$T_{сл}$, ме
1	2	3	4	5	6	7	8	9	10	11	12	13
ГОЭС-4	1.047	1.745	0.01	0.3	1.885	0.628	314	6.0	209	20.0	11	170
Туре239	2.094	17.453	0.01	0.3	0.377	0.251	314	1.2	209	4.0	11	34

В таблице 3 (см. столбцы 4 и 5) заданы исходные данные моделей: погрешность восстановления $\delta = 0.01$ для трёх моделей, рассогласование параметров $k = 0.3$ для составной полуволны. В столбцах 8÷13 расположены результаты моделирования по выражениям (21), (22), (26), из которых следует, что ОЭС Туре239 по величине интервала реверса быстрее ОЭС ГОЭС-4 в пять раз и имеет угол останова меньше в 2.5 раза. Это обеспечивает большую оперативность при переключении в режиме сопровождения с одной цели на другую и определяет:

Свойство 15. Модели гармонических полуволн и составных полуволн определяются двумя интегральными характеристиками ОЭС интервалом реверса и углом «мгновенная» останова, которые можно определять через паспортные параметры угловую скорость и ускорение для оценки оперативности реакции ОЭС на смену цели.

Свойство 16. Коэффициент Q превышения величины частоты гармонической полуволны и составных полуволн величинами частот равномерной дискретизации зависит от типа полуволны (гармоническая полуволна или составная полуволна), типа равномерной аппроксимации (ступенька или отрезок прямой), величин максимальной, относительной погрешности аппроксимации.

В процессе эксплуатации ОЭС может произойти разряд аккумулятора, питающего ОЭС, что приведёт к снижению величины максимальной скорости в соответствии с (11). Примем $k = 0.3$ в таблице 3 (см. столбцы 5, 10 и 11). В этом случае траектория ОЭС в режиме реверса соответствует выражению (22) или (23). Это, в обоих случаях (см. рисунок 7), приводит к увеличению интервала дискретизации времени $T_{лс}$ или $T_{гс}$ по сравнению с интервалом $T_{сс}$. Данной ситуации соответствует свойство подобное Свойству 11, но с возможностью задания необходимой погрешности восстановления сигнала по выборкам:

Свойство 17. Уменьшение величины коэффициента k согласования параметров сигнала (управляемого технологического процесса) приводит к увеличению интервала дискретизации времени $T_{лс} = T_{л}/Q_{лс}$ или $T_{гс} = T_{г}/Q_{гс}$ по сравнению с $T_{сс} = T_{с}/Q_{сс}$ при задании одинаковой величины погрешности аппроксимации $\delta_{лс} = \delta_{гс}$ или приводит к обоснованному увеличению единицы измерения системного времени.

Свойство 14 подтверждается количественно результатами в таблице 3 (см. столбцы 8, 9 и 12, 13). Применение в ЦСП линейной аппроксимации выборок из сигналов с датчика азимута ОЭС в подсистеме сбора информации СРВ вместо ступенчатой аппроксимации сокращает количество выборок в $Q_{сс}/Q_{сл} = 28.54$ раза (при одинаковой точности восстановления сигнала $\delta_{сс} = \delta_{сл} = 0.01$) и, соответственно, увеличивает интервал дискретизации в $T_{сл}/T_{сс} = 28.33$ раза.

7. Заключение. В технических системах существуют сигналы и процессы, характеризующиеся трудностями определения спектра (СТОС), например, от объектов управления робастных автоматизированных систем управления реального времени (СРВ):

- процесс вывода пароварочной камеры на режим варки колбас связан с сигналами температуры и влажности пара, которые необходимо наиболее быстро привести к показателям, оговоренным в соответствующих технических условиях, и в процессе варки регулировать;
- в оптико-электронных системах (ОЭС) необходимо управлять углом азимута и углом места вращения оптических и лазерных датчиков и считывать их значения с необходимой точностью;

– на высокоскоростных магистралях (ВСМ) необходимо регулировать скорость состава и предсказывать его положение при экстренном торможении, если необходимо обеспечить безопасность движения составов с интервалом в три минуты со скоростью 350-400 км/ч...

Примеры имеют общие свойства физических параметров сигналов или процессов: максимальную скорость и максимальное ускорение изменения, максимальное отклонение от установившегося значения. Замерим максимальные значения этих параметров для пароварочной камеры, ОЭС, локомотива и участка пути или другими словами проведём обследование объекта управления на стадии проектирования системы управления реального времени.

Используем максимальные значения параметров: определим условие их согласования в соответствии с предложенными новыми моделями полуволн и их свойствами. Выберем гармоническую полуволну, если максимальные параметры согласованы или составную полуволну, если они рассогласованные. Новые модели этих полуволн позволяют по математическим выражениям статьи определить эквивалентную, согласованную частоту выборок f_c из сигнала или процесса СТОС. Эта частота f_c гарантированно обеспечит заданную в техническом задании на СРВ погрешность восстановления сигнала или процесса по выборкам из него, если аппроксимацию проводить ступеньками или отрезками прямых линий. В новых моделях полуволн частота выборок f_c обоснованно ($f_c = 2Qf$) связана с частотой среза f цифрового ФНЧ и необходима для проектирования процесса цифровой предварительной обработки сигналов в цифровом процессоре сигналов (ЦПС). После восстановления в ЦПС множество сигналов поступает и совместно обрабатывается в центральном процессоре системы реального времени.

Аналогично, если аппарат вейвлетов (wavelet) позволяет исследовать и интерпретировать представления сигналов в спектральной области, то предложенный в статье аппарат составных полуволн позволяет интерпретировать сигналы и процессы СТОС во времени.

Исследование моделей составных полуволн показывает, что уменьшение величины коэффициента согласования параметров сигнала (управляемого технологического процесса) приводит к увеличению интервала дискретизации времени (более чем на порядок) или приводит к обоснованному увеличению единицы измерения системного времени и тем самым к снижению загрузки ЦПС.

Литература

1. Никифоров В.В., Баранов С.Н. Статическая проверка корректности разделения ресурсов в системах реального времени // Труды СПИИРАН. 2017. №3(52). С.137-156.
2. Ровелли К. Срок времени / Пер. с итал. Д. Баюка // М.: Издательство АСТ: CORPUS. 2020. 224 с.
3. Arpe S.D. Signals and systems: Principles and applications // Cambridge: Cambridge University Press. 2016. 768 p.
4. Лазарев В.Л. Робастное управление в биотехнологической промышленности // СПб.: Университет ИТМО; ИХиБТ. 2015. 196 с.
5. Контрольно-измерительное оборудование R&S 2019.– Каталоги Rohde&Schwarz, 2019. 256 с. https://www.Rohdeschwarz.com/ru/top_topics_russia/catalogs/catalogs_231625.html (дата обращения: 12.03.2019).
6. Якимов Е.В. Цифровая обработка сигналов: учебное пособие // Томск: Издательство ТПУ. 2011. 168 с.
7. Кестер У., Брайэнт Д. Проектирование систем цифровой и смешанной обработок сигналов. / Пер. под ред. А.А. Власенко // М.: Техносфера. 2010. 384с.
8. Сенкевич Ю.И., Марапулец Ю.В., Луковенкова О.О., Солодчук А.А. Методика выделения информативных признаков в сигналах геоакустической эмиссии // Тр. СПИИРАН. 2019. №5(18). С.1066–1092.
9. Петренко В.И. и др. Прогнозная оценка траектории руки оператора для решения обратной задачи динамики при копирующем управлении // Труды СПИИРАН. 2019. №1(18). С. 123–147.
10. Demydyuk M.V., Hoshovs'ka N. Parametric optimization of the transport operations of a two-link manipulator // J. Math. Sci. 2019. vol. 238. no. 2. pp.174–188.
11. Manyam S.G. Optimal dubins paths to intercept a moving target on a circle. // Proc. Am. Control Conf. 2019. (July 2019). pp.828–834.
12. Poulsen D.R., Davis J.M., Gravagne I.A. Optimal Control on Stochastic Time Scales. IFAC-PapersOnLine. 2017. Vol. 50. no. 1. pp. 14861–14866.
13. Майоров Б.Г. Способ непрерывной оптической связи с низколетающей целью // Патент RU №2 715 499 С1. 28.02.2020.
14. Vavilov S.A., Lytaev M.S. Calibration and verification of models defining radar visibility zones in marine geoinformation systems // Proceedings of the 8th international Symposium on Information Fusion and Intelligent Geographic Information Systems (IF&IGIS'17). 2018. pp. 115–125.
15. Кузнецов О.П., Базенков Н.И., Болдышев Б.А. Асинхронная дискретная модель химических взаимодействий в простых нейронных системах // Искусственный интеллект и принятие решений. 2018. № 2. С. 3–20.
16. Kar N.R. Production and Applications of Radiopharmaceuticals: A Review // Int. J. Pharm. Investigation. 2019. vol. 9. no. 2. pp. 36–42.
17. Фельдбаум А.А. Оптимальные процессы в системах автоматического регулирования // Автоматика и телемеханика. 1953. № 6(14). С. 712–728.
18. Uneyama T., Miyaguchi T., Akimoto T. Relaxation Functions of the Ornstein Uhlenbeck Process with Fluctuating Diffusivity. Physical Review. 2019. vol. 99. no. 3. pp. 21–27.
19. Хлистунов В.Н. О погрешности аппроксимации дискретных методов измерения // Приборостроение. 1960. №5. С. 3–5.
20. Майоров Б.Г. Обобщенный критерий наибольшего отклонения входных сигналов систем управления // Автоматика и Телемеханика. 2005. №10. С. 148–155; англ. пер.: B.G. Maierov. The generalized maximum deviation criterion for input signals of control systems. //Autom. Remote Control. 66:10 (2005). pp. 1666–1672.

21. Шеннон К. Работы по теории информации и кибернетике / Пер. с англ. // М.: Издательство иностранной литературы. 1963. 830 с.
22. Котельников В.А. О пропускной способности «эфира» и проволоки в электросвязи // Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности. М.: Редакция управления связи РККА. 1933. С. 4–5.
23. Nyquist H. Certain topics in telegraph transmission theory // Trans. AIEE. 1928. vol. 47. iss. 2. pp. 617–644.
24. Хлистунов В.Н. О применении теоремы Котельникова к дискретной измерительной технике // Измерительная техника. 1961. № 3. С. 15–28
25. Майоров Б.Г. Исследование теоремы Котельникова и применение результатов для определения величины дискретизации по времени входных сигналов систем управления // Высокопроизводительные вычислительные системы и микропроцессоры // Сб. научных трудов ИМВС РАН. 2004. №7. С. 76–82.
26. Майоров Б.Г. Восстановление сигнала в системах реального времени по равномерным выборкам с уменьшением интервала Найквиста // Системы и средства информатики. 2019. №2(29). С. 95–112.
27. Табачук И.С., Ташкеев Л.Л. Угрозы с предельно малых высот // Воздушно-космическая оборона. 2007. №1. С. 50–57.

Майоров Борис Геннадьевич — Ph.D., ведущий научный сотрудник, центр специальных систем, АО "Научно-производственное предприятие "Рубин". Область научных интересов: цифровая обработка сигналов, системы сбора и отображения информации в реальном времени. Число научных публикаций — 86. bgmayorov@yandex.ru; ул. Байдукова, 2, 440015, Пенза, Россия; р.т.: +7(8412)204777.

B. MAYOROV

**PROPERTIES OF HARMONIC AND COMPOSITE HALF-WAVES,
DETERMINATION OF THE UNIFORM TIME SAMPLING
INTERVAL OF DIGITAL SIGNAL PROCESSORS**

Mayorov B. Properties of Harmonic and Composite Half-Waves, Determination of the Uniform Time Sampling Interval of Digital Signal Processors.

Abstract. When building autonomous real-time systems (RTS), it is necessary to solve the problem of optimal multitasking loading of a number of parallel functioning digital signal processors. One of the reserves for achieving the desired result is the implementation of samples from the sensor signals of information about the magnitude of the signal most rarely in time. In this case, it is necessary to provide a linear or stepwise approximation of the signal by samples with an acceptable reconstruction error. One of the system tasks of these processors is filtering signals or limiting the spectrum to the cutoff frequency. A distinctive feature of the approach proposed in the article is the fulfillment of the condition: if the measurement of this frequency is difficult (for example, in the electromechanical means of the RTS), then for such signals it is proposed to match the maximum values of the harmonic half-wave parameters: approximation error, speed and acceleration. The study opens up the prospect of applying new approaches to sampling the time of signals in the amplitude-time domain and determining the equivalent cutoff frequency of the signal spectrum for such signals. In this article, the dependences of the value of the unit of system time for input-output of data on the degree of agreement between the maximum values of the signal parameters are obtained. A mathematical model of the extreme behavior of a signal between two adjacent samples is given in the form of a harmonic half-wave. The study is also extended to convex composite harmonic functions, according to which the signal can deviate from the results of a linear or stepwise approximation of the signal for these samples. The comparison of the models by the value of the relative time sampling intervals, depending on the degree of matching of the maximum parameters of the harmonic half-wave, is carried out. When comparing, in addition to these maximum parameters, the relationship of the maximum signal speed with the error of approximating the samples by steps and the relationship of the maximum acceleration of the signal with the maximum error of the linear approximation was taken into account. The results make it possible to determine the duration of the intervals of uniform sampling of the signal time based on the results of the inspection of the control object, substantiate a significant increase in the sampling interval of time or a similar increase in the number of tasks to be solved per unit of system time.

Keywords: real-time control systems, digital signal processors, signal filtering, signal model - harmonic half-wave, composite half-waves, uniform samples, uniform time sampling, stepwise and linear signal approximation, admissible error.

Mayorov Boris — Ph.D., Leading researcher, Center for special systems, Joint Stock Company Scientific and Production Enterprise Rubin. Research interests: digital signal processing, systems for collecting and displaying information in real-time. The number of publications — 86. bgmayorov@yandex.ru; 2, Baidukova St., 440015, Penza, Russia; office phone: +7(8412)204777.

References

1. Nikiforov V.V., Baranov S.N. [Static verification of the correctness of resource sharing in real-time systems]. Trudy SPIIRAN – SPIIRAS Proceedings. 2017. vol. 3(52). pp.137-156. (In Russ.).
2. Rovelli K. Srok vremena [Time period]. Translation from italian D. Bajuka. Moscow: Publishing house ACT: CORPUS, 2020. 224 p. (In Russ.).
3. Apte S.D. Signals and systems: Principles and applications. Cambridge: Cambridge University Press. 2016. 768 p.
4. Lazarev V.L. Robastnoe upravlenie v biotekhnologicheskoy promyshlennosti [Robust control in the biotechnological industry]. SPb.: ITMO University; IHiBT. 2015. 196 p. (In Russ.).
5. Kontrol'no-izmeritel'noe oborudovanie R&S 2019 – Katalogi Rohde&Schwarz. [Test and measurement equipment R&S 2019.– Catalogs Rohde& Schwarz]. 2019. 256 p. Available at: www.rohdeschwarz.com/ru/top_topics_russia/catalogs/catalogs_231625.html (accessed: 12.03.2019). (In Russ.).
6. Yakimov E.V. Zifrovaya obrabotka signalov: uchebnoe posobie. [Digital signal processing: a tutorial]. Tomsk: TPU publishing house. 2011.168 p. (In Russ.).
7. Kester W., Bryant D. Proektirovanie system zifrovoy i smeshannoy obrabotki signalov. [Design of digital and mixed signal processing systems]. Translation edited by A.A. Vlasenko. Moscow: Technosphere. 2010. 384 p. (In Russ.).
8. Senkevich Yu.I., Marapulez Yu.V., Lukovenkova O.O., Solodchuk A.A. [Technique for identifying informative features in geoacoustic emission signals]. Trudy SPIIRAN – SPIIRAS Proceedings. 2019. vol. 5(18). pp. 1066–1092. (In Russ.).
9. Petrenko V.I. et. al. [Predictive estimate of the trajectory of the operator's hand for solving the problem of dynamics with copying control]. Trudy SPIIRAN – SPIIRAS Proceedings. 2019. vol. 1(18). pp. 123–147. (In Russ.).
10. Demydyuk M.V., Hoshovs'ka N. Parametric optimization of the transport operations of a two-link manipulator. J. Math. Sci. 2019. vol. 238. no. 2. pp. 174–188.
11. Manyam S.G. Optimal dubins paths to intercept a moving target on a circle. Proc. Am. Control Conf. 2019. (July 2019). pp. 828–834.
12. Poulsen D.R., Davis J.M., Gravagne I.A. Optimal Control on Stochastic Time Scales. IFAC-PapersOnLine. 2017. vol. 50. no. 1. pp. 14861–14866.
13. Mayorov B.G. [Method of continuous optical communication with low-flying target] Patent RU. no. 2 715 499 C1. 02.28.2020. (In Russ.).
14. Vavilov S.A., Lytaev M.S. Calibration and verification of models defining radar visibility zones in marine geoinformation systems. Proceedings of the 8th international Symposium on Information Fusion and Intelligent Geographic Information Systems (IF&IGIS'17). 2018. pp. 115–125.
15. Kuznetsov O.P., Bazenkov N.I., Boldyshev B.A. [Asynchronous discrete model of chemical interactions in simple neural systems] *Iskusstvennyj intellekt i prinyatie reshenij – Artificial Intelligence and Decision Making*. 2018. no. 2. pp. 3–20.
16. Kar N.R. Production and Applications of Radiopharmaceuticals: A Review. Int. J. Pharm. Investigation. 2019. vol. 9. no. 2. pp. 36–42.
17. Feldbaum A.A. [Optimal processes in automatic control systems]. *Avtomatika i telemekhanika. [Automation and telemechanics]*. 1953. no. 6(14). pp. 712–728. (In Russ.).
18. Uneyama T., Miyaguchi T., Akimoto T. Relaxation Functions of the Ornstein Uhlenbeck Process with Fluctuating Diffusivity. *Physical Review*. 2019. vol. 99. no. 3. pp. 21–27.
19. Khlistunov V.N. [On the error of approximation of discrete measurement methods] *Priborostroenie – Instrumentation*. 1960. no. 5. pp. 3–5. (In Russ.).

20. Mayorov B.G. [Generalized criterion for the greatest deviation of input signals of control systems]. *Avtomatika i telemekhanika – Automation and telemechanics*. 2005. no. 10. pp.148–155. (In Russ.). trans. eng.: B.G. Maiorov. The generalized maximum deviation criterion for input signals of control systems. // *Autom. Remote Control*. 66:10 (2005). pp. 1666–1672.
21. Shannon C. *Raboty po teorii informatsii i kibernetike. Raboty iz razlichnykh zhurnalov* [Works on the theory of information and cybernetics. Works from different magazines]. Transfer from English. Moscow. Publishing house of foreign literature, 1963. 830 p. (In Russ.).
22. Kotelnikov V.A. [About the flow capacity of "broadcast" and wire in telecommunication]. *Materialy k I Vsesoyuznomu siesdu po voprosam tekhnicheskoy rekonstruktsii dela svyazi i rasviniya slabotochnoy promyshlennosti*. [Materials for the 1st All-Union Congress on the issues of technical reconstruction of communications and the development of low-current industry.] Published by the editorial board of the communications department of the RKKA. 1933. Moscow. pp. 4–5. (In Russ.).
23. Nyquist H. Certain topics in telegraph transmission theory. *Trans. AIEE*. 1928. vol. 47. iss. 2. pp. 617–644.
24. Khlistunov V.N. [About application of the theorem of Kotelnikov to discrete measuring technique]. *Izmeritel'naya tekhnika – Measuring technique*. 1961. no. 3. pp. 15–28. (In Russ.).
25. Mayorov B.G. [Investigation of Kotelnikov's theorem and application of the results to determine the value of time sampling of input signals of control systems]. *Visokoproizvoditel'nye vychislitel'nye sistemy i mikroprozessory: Sbornik nauchnykh trudov IMVS RAN – High-performance computing systems and microprocessors: Collection of articles scientific works of IMVS RAS*. Moscow. 2004. no. 7. pp. 76–82. (In Russ.).
26. Mayorov B.G. [Reconstruction of a signal in real-time systems from uniform samples with a decrease in the Nyquist interval]. *Sistemy i sredstva informatiki – Systems and means of informatics*. 2019. no. 2 (29). pp. 95–112. (In Russ.).
27. Tabachuk I.S., Tashkeyev L.L. [Threats from extremely low altitudes]. *Vozdushno-kosmicheskaya oborona – Aerospace defense*. 2007. no. 1. pp. 50–57. (In Russ.).

В.Л. ЯКИМОВ, Г.Н. МАЛЬЦЕВ
**ГИБРИДНЫЕ СЕТЕВЫЕ СТРУКТУРЫ И ИХ
ИСПОЛЬЗОВАНИЕ ПРИ ДИАГНОСТИРОВАНИИ СЛОЖНЫХ
ТЕХНИЧЕСКИХ СИСТЕМ**

Якимов В.Л., Мальцев Г.Н. Гибридные сетевые структуры и их использование при диагностировании сложных технических систем.

Аннотация. Предложен подход к техническому диагностированию сложных технических систем по результатам обработки телеметрической информации внешней системой контроля и диагностирования с использованием гибридных сетевых структур. Рассмотрен принцип построения диагностических комплексов сложных технических систем, обеспечивающий автоматизацию процесса технического диагностирования и основанный на использовании при обработке телеметрической информации моделей в виде гибридных сетевых структур, включающих многослойные нейронные сети и дискретные байесовские сети со стохастическим обучением. Разработаны модель изменения параметров технического состояния сложных технических систем на основе многослойных нейронных сетей, позволяющая сформировать вероятностную оценку отнесения текущей ситуации функционирования сложной технической системы к множеству рассмотренных ситуаций функционирования по отдельным телеметрируемым параметрам, и многоуровневая иерархическая модель технического диагностирования сложных технических систем на основе дискретной байесовской сети со стохастическим обучением, позволяющая агрегировать полученную от нейросетевых моделей информацию и распознавать текущую ситуацию функционирования сложной технической системы. В условиях нештатных ситуаций функционирования сложной технической системы по результатам обработки телеметрической информации локализуются неисправные функциональные узлы и формируется объяснение причины возникновения нештатной ситуации. Детализированы этапы реализации технического диагностирования сложных технических систем с использованием предложенных гибридных сетевых структур при обработке телеметрической информации. Представлен пример использования разработанного подхода к решению задач технического диагностирования бортовой системы космического аппарата. Показаны преимущества предлагаемого подхода к техническому диагностированию сложных технических систем в сравнении с традиционным подходом, основанном на анализе принадлежности значений телеметрируемых параметров заданным допускам.

Ключевые слова: сложная техническая система, техническое диагностирование, гибридная сетевая структура, байесовская сеть, стохастическое обучение.

1. Введение. В различных сферах деятельности в процессе эксплуатации сложных технических систем (СТС) широко используется принцип ситуационного управления [1]. При этом принципе управление СТС осуществляется с учетом текущей ситуации функционирования, определяемой условиями применения СТС, техническим состоянием и функциональными возможностями. Принимаемые решения по управлению и применению СТС по назначению в значительной степени определяются результатами технического диагностирования, характеризующими текущее

техническое и функциональное состояния СТС [2, 3]. Отсюда следует важность задачи технического диагностирования СТС, дающего необходимую информацию для обоснованного принятия решений по управлению и применению СТС.

Особенностью процесса технического диагностирования современных СТС является распределение задач диагностирования между встроенными и внешними средствами контроля и диагностирования. Также необходимо наличие баз данных с набором типовых ситуаций функционирования СТС и их признаков, формируемых на основе принимаемой телеметрической информации (ТМИ). Это характерно, в частности, для технического диагностирования находящихся на орбите космических аппаратов (КА) [2–6]. Наличие баз данных ситуаций функционирования СТС позволяет выявлять и устранять (парировать) рассмотренные в них нештатные ситуации с помощью встроенных средств контроля, технического диагностирования и восстановления работоспособного состояния СТС. В то же время при возникновении нештатных ситуаций, не рассмотренных в базах данных СТС, распознавание текущей ситуации функционирования СТС осуществляется внешней системой контроля технического состояния и диагностирования СТС по результатам обработки ТМИ. Восстановление ее работоспособного состояния также осуществляется внешней системой управления.

С использованием внешних средств технического диагностирования СТС выявляются отказы и нештатные ситуации, отсутствующие в базе данных, а также ситуации функционирования, характеризующиеся существенной априорной неопределенностью значений временных рядов телеметрируемых параметров (ТМП) и сложностью выявления причины их возникновения, необходимостью проведения углубленного анализа всей имеющейся диагностической информации. У разработчиков КА подобного рода ситуации функционирования получили название «тяжелых» отказов [4, 6]. Возможности распознавания таких ситуаций функционирования СТС на основе традиционно используемого при контроле их технического состояния допускового контроля ограничены. Различные ситуации функционирования СТС допуска на значения отдельных ТМП могут пересекаться. Отличительным признаком подобных ситуаций функционирования СТС может быть характер изменения отдельных ТМП, а не область определения их значений. Поэтому в основу технического диагностирования СТС, для которых характерны такие ситуации функционирования, должно быть положено диагностирование с использованием моделей изменения их ТМП. При

этом допуски могут быть использованы для контроля работоспособности СТС и задаваться как для отдельных ТМП, так и для обобщенных параметров технического состояния, определяемых по совокупности ТМП.

В настоящее время для большинства СТС задачи углубленного анализа полных потоков ТМИ решаются в отложенном режиме в течение достаточно продолжительного периода времени. Так, при техническом диагностировании находящихся на орбите КА для углубленного анализа таких ситуаций функционирования используются принятые с КА полные потоки ТМИ. Привлекаются эксперты и разработчики бортовой аппаратуры. Космический аппарат на время анализа технического состояния и нештатной ситуации функционирования бортовой аппаратуры переводится в специальный дежурный режим. Поэтому автоматизация процесса обработки полных потоков ТМИ о техническом состоянии СТС и повышение оперативности проведения углубленного анализа ситуаций их функционирования имеет важное значение для совершенствования диагностических комплексов СТС. Она позволяет оперативно выявлять и парировать нештатные ситуации и реализовать перспективные схемы ситуационного и проактивного управления применением СТС [7–10].

В статье рассмотрен подход к техническому диагностированию СТС по результатам обработки ТМИ внешней системой контроля и диагностирования с использованием моделей в виде гибридных сетевых структур, включающих многослойные нейронные сети и дискретные байесовские сети со стохастическим обучением. Многослойные нейронные сети, реализующие вывод о ситуациях функционирования аппаратных модулей и подсистем СТС по отдельным ТМП, являются иерархически связанными между собой моделями изменения и обработки ТМИ. Дискретные байесовские сети со стохастическим обучением, реализующие вывод о ситуациях функционирования СТС в целом, являются моделью распознавания текущей ситуации функционирования СТС.

2. Общая характеристика подхода к техническому диагностированию СТС по результатам обработки ТМИ с использованием моделей в виде гибридных сетевых структур. Предлагаемый подход реализует идеи технического диагностирования СТС. Он основан на моделях, и их многомодельном описании. Этот подход в настоящее время получает широкое распространение при решении задач технического диагностирования и управления СТС [11–14]. При этом для описания процессов функционирования и

технического диагностирования СТС используется широкий спектр математических моделей: вероятностные модели, производственные модели, модели на основе различных автоматов, сетей и другие модели [12, 15–20]. Выбранные типы моделей – многослойные нейронные сети и дискретные байесовские сети – хорошо согласуются с задачей технического диагностирования СТС. Эти СТС имеют многоуровневую иерархическую структуру и сложный характер зависимости между параметрами, отражающие техническое состояние и функциональные возможности систем. Модели их объединяются в агрегированные гибридные сетевые структуры.

Применение моделей изменения ТМП СТС на основе многослойных нейронных сетей позволяет обеспечить: высокую достоверность распознавания ситуаций функционирования аппаратных модулей и подсистем СТС в условиях значительной неопределенности значений и существенно нелинейного характера ТМП; возможность обработки разнородной (не только телеметрической) информации о ситуациях функционирования СТС; распараллеливание процессов обработки полных потоков ТМИ для достижения высокой оперативности ее углубленного анализа [21-24]. Модели диагностирования СТС на основе дискретной байесовской сети позволяет объяснить полученные результаты распознавания ситуаций функционирования СТС и определить наиболее вероятные причины возникновения ранее не рассмотренных нештатных ситуаций, что очень важно для формирования баз данных систем поддержки принятия решения по управлению СТС [25].

Существенной особенностью рассматриваемых гибридных моделей в виде сетевых структур с точки зрения их использования при техническом диагностировании СТС является возможность решать в рамках одной агрегированной модели не только задачи контроля технического состояния и диагностирования СТС, но и задачу выявления негативных тенденций изменения технического состояния СТС. Она в отличие от задачи экстраполяции временных рядов ТМП СТС, является востребованной на практике, но достаточно слабо освещена в литературе [33]. При этом предлагаемая реализация дискретной байесовской сети имеет отличия от общепринятых реализаций, рассматриваемых в большинстве источников [26-32]. Она предполагает совместную обработку двух видов информации об одних и тех же ситуациях функционирования СТС, получаемую на основе знаний экспертов и по результатам обработки ТМИ. Это нашло отражение в ее архитектуре и функционировании и потребовало разработки оригинального стохастического подхода к обучению сети

на основе дискретно-событийного имитационного моделирования. В доступных источниках примеры реализации таких сетей и подробное описание их функционирования практически отсутствуют, а подходы к согласованию разнородных данных в байесовских сетях рассмотрены в незначительной степени [30].

Подход к техническому диагностированию СТС по результатам обработки ТМИ с применением моделей в виде гибридных сетевых структур основан на следующей процедуре. Это преобразование исходных значений разнородных ТМП во множество однородных вторичных признаков с использованием множества нелинейных моделей $\{M_{v,p}\}$, характеризующих изменение различных ТМП в различных ситуациях функционирования СТС. Признаками технического состояния СТС являются вероятности отнесения текущей динамической реализации v -ых ТМП в p -ых ситуациях функционирования к моделям $\{M_{v,p}\}$. Данные вероятности, как признаки технических состояний, связываются между собой в рамках агрегированной модели M_{Σ} , отражающей многоуровневые иерархические взаимосвязи параметров СТС. Результатом технического диагностирования является распознавание ситуаций функционирования СТС. Модели $\{M_{v,p}\}$ и M_{Σ} создаются на этапе проектирования СТС и корректируются на последующих этапах их жизненного цикла. Так, нештатная ситуация, выявленная внешней системой контроля и диагностирования и отработанная внешней системой управления, добавляется в базу данных ситуаций функционирования СТС.

Реализация принципов технического диагностирования СТС, основанного на моделях, и их многомодельное описание являются наиболее продуктивными при автоматизированном решении задач обработки полных потоков ТМИ о техническом состоянии СТС в условиях априорной неопределенности. Как показывает практика, включение нелинейных моделей в механизмы определения технического состояния анализируемой СТС и распознавания нештатных ситуаций ее функционирования позволяет получать правильные решения в сложных условиях функционирования СТС, когда алгоритмы только на основе логических правил и допусков недостаточно эффективны [11].

Достоинством предлагаемого подхода к построению моделей технического диагностирования СТС с использованием гибридных сетевых структур является возможность повышения достоверности решения задачи распознавания текущей ситуации функционирования СТС. Это делается за счет:

- формирования пространства однородных вторичных признаков (в том числе нелинейных), в котором задача обучения модели распознавания ситуации функционирования СТС будет решаться быстрее. При заданном времени обучения будет достигаться более высокая достоверность их распознавания;

- расширения пространства диагностических признаков при увеличении количества ситуаций функционирования СТС на каждом уровне агрегированной модели M_{Σ} ;

- обеспечения инвариантности вторичных признаков к режимам работы и условиям функционирования СТС;

- возможности достижения низкой чувствительности вторичных признаков к различного рода возмущениям во временных рядах ТМП за счет их фильтрации с использованием моделей $\{M_{v,p}\}$.

Гибридный характер используемых моделей технического диагностирования СТС определяется объединением в них моделей двух типов: многослойных нейронных сетей и дискретных байесовских сетей. Сетевой характер используемых моделей обусловлен многоуровневой иерархической структурой СТС и присущей им структурной сложностью на каждом уровне представления. При этом сами гибридные сетевые модели имеют два уровня описания ситуаций функционирования СТС: нижний уровень для ситуаций функционирования аппаратных модулей и подсистем СТС по отдельным ТМП и верхний уровень для ситуаций функционирования СТС в целом по всему анализируемому потоку ТМИ. Выполним формализованное описание гибридных сетевых моделей технического диагностирования СТС, начиная с верхнего уровня, на котором осуществляется принятие решения о техническом состоянии СТС и распознавание текущей ситуации ее функционирования.

3. Формализованное описание процесса технического диагностирования СТС по результатам обработки ТМИ с использованием дискретных байесовских сетей. В общем случае агрегирование информации от динамических моделей изменения ТМП $\{M_{v,p}\}$ при оценке технического состояния СТС может быть реализовано различными способами [13–21]. Математические модели дискретных байесовских сетей [11, 24–30] позволяют:

- реализовать агрегированную модель M_{Σ} технического диагностирования СТС и распознавания ее текущей ситуации функционирования по результатам обработки совокупности ТМП, образующих полный поток ТМИ;

- насытить модель M_{Σ} имеющимися экспертными знаниями о

функционировании СТС и объяснить полученные результаты распознавания ситуаций функционирования СТС.

Важными достоинствами дискретных байесовских сетей при решении задач технического диагностирования СТС являются математически доказанные свойства непротиворечивости вывода, возможность его в условиях неполного вектора диагностических признаков, осуществимость работы с задачами большой размерности, многообразие подходов к их обучению и применению [11, 21–30]. Кроме того, использование диагностических моделей на основе дискретных байесовских сетей позволяет выполнить требование к автоматизации процессов обработки ТМИ и технического диагностирования СТС в интересах поддержки принятия решений по управлению СТС. В соответствии с ними по результатам распознавания ситуаций функционирования СТС требуется формирование вариантов принятия решений по управлению СТС и предъявление лицу, принимающему решение, объяснения полученных результатов [11]. Именно поэтому использование дискретных байесовских сетей является одним из наиболее адекватных подходов для реализации диагностических моделей СТС, позволяющим реализовать логико-вероятностные рассуждения и сформировать на основе базы данных СТС обоснованный вывод о текущей ситуации ее функционирования.

Структурная схема модели распознавания ситуаций функционирования СТС на основе дискретной байесовской сети, реализующая рассматриваемый подход к техническому диагностированию СТС, представлена на рисунке 1. Модель отражает иерархическую структуру СТС и пути развития ситуаций функционирования между уровнями ее иерархии. Модель представляет собой направленный граф, без циклов. Дуги графа отображают причинно-следственные связи между переменными, благодаря чему дискретные байесовские сети также называют причинно-следственными сетями. Структура предлагаемой многоуровневой иерархической модели M_{Σ} содержит пять уровней иерархии: отдельных ТМП, аппаратных модулей, подсистем, функциональных систем и СТС в целом. Такая декомпозиция структуры СТС позволяет описать широкий класс СТС различного назначения.

Количество уровней иерархии модели M_{Σ} было выбрано на основе изучения типовых структур СТС и известных подходов к такому делению [3], а также анализа сообщений о неисправностях в СТС различного назначения. Для описания процесса принятия

решения по управлению СТС дискретная байесовская сеть может быть дополнена шестым уровнем с вершинами, характеризующими обобщенные состояния СТС. В качестве них выступают состояния: штатное функционирование СТС, нештатные ситуации (частично-работоспособное состояние СТС и др.) [2, 4, 6]. Количество уровней иерархии может варьироваться, в зависимости от требуемой степени детализации при локализации неисправных функциональных узлов СТС.

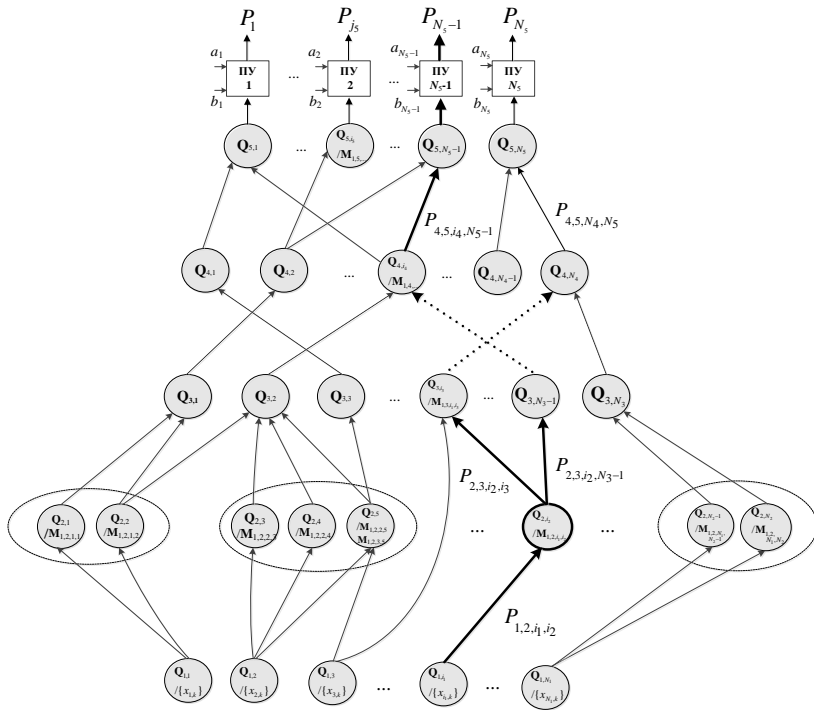


Рис. 1. Структурная схема модели распознавания ситуаций функционирования СТС на основе дискретной байесовской сети

Множество значений всех учитываемых при техническом диагностировании ТМП определяет множество начальных состояний модели. На каждом последующем уровне (аппаратных модулей, подсистем и функциональных систем) имеет место конечное число типовых штатных ситуаций функционирования и нештатных ситуаций. Выходным результатом является одна из возможных

ситуаций функционирования на уровне СТС, определяющих множество поглощающих состояний модели, число которых может быть значительным и составлять несколько сотен ситуаций.

Введем следующие обозначения:

N_l – количество ситуаций функционирования на каждом l -ом уровне (количество вершин), где $l=1\dots 5$ – номер уровня;

N_1 – количество начальных состояний;

N_5 – количество поглощающих состояний;

$\{Q_{l,i} | \{x_{l,i}\}\}$ – множество вершин $\{Q_{l,i}\}$ уровня дискретной байесовской сети с номером $l=1$, соответствующих динамическим реализациям i -ых ТМП и представляющих собой множество значений $\{x_{l,i}\}$;

$\{Q_{l,i}\}$ – множество вершин уровня дискретной байесовской сети с номером l , характеризующих штатные ситуации функционирования и нештатные ситуации в элементах СТС различного уровня иерархии ($l=2,3,4$) и СТС в целом ($l=5$);

$\{P_{l-1,l}\}$ – множество матриц вероятностных переходов между смежными уровнями, начиная со второго, характеризующих причинно-следственные связи между ситуациями функционирования;

$\{P_{1 \times l}^*\}$ – множество матриц переходов между первым уровнем и остальными уровнями, начиная со второго, отражающих условные вероятности возникновения ситуаций функционирования по результатам соотнесения значений ТМП и нейросетевых динамических моделей их изменения $M_{1,l,i,i}$;

P_l^* – вектора вероятностей достижения вершин каждого уровня с номерами $l=1\dots 5$.

Поглощающие вершины на пятом уровне дискретной байесовской сети соответствуют множеству рассмотренных ситуаций функционирования СТС, которое, в общем случае, включает штатные ситуации функционирования СТС и рассмотренные нештатные ситуации из базы данных СТС.

Первый уровень модели связан со всеми остальными уровнями через разреженные матрицы вероятностей переходов $P_{1 \times l}^*$:

$$P_{1 \times l}^* = \begin{pmatrix} P_{1,l,1,1} & P_{1,l,1,2} & \dots & P_{1,l,1,N_l} \\ P_{1,l,2,1} & P_{1,l,2,2} & \dots & P_{1,l,2,N_l} \\ \dots & \dots & \dots & \dots \\ P_{1,l,N_1} & P_{1,l,N_1,2} & \dots & P_{1,l,N_1,N_l} \end{pmatrix}, \quad 2 \leq l \leq 5. \quad (1)$$

Элементы матрицы вероятностей переходов $\mathbf{P}_{l \times l}^*$ вида (1) определяются по результатам обработки ТМП с использованием динамических моделей их изменения \mathbf{M}_{1,l,i_j,i_j} .

Смежные уровни, начиная со второго, связаны между собой разреженными матрицами вероятностей переходов $\mathbf{P}_{l-1 \times l}$:

$$\mathbf{P}_{l-1 \times l} = \begin{pmatrix} P_{l-1,l,1,1} & P_{l-1,l,1,2} & \dots & P_{l-1,l,1,N_l} \\ P_{l-1,l,2,1} & P_{l-1,l,2,2} & \dots & P_{l-1,l,2,N_l} \\ \dots & \dots & \dots & \dots \\ P_{l-1,l,N_{l-1},1} & P_{l-1,l,N_{l-1},2} & \dots & P_{l-1,l,N_{l-1},N_l} \end{pmatrix}, 3 \leq l \leq 5. \quad (2)$$

Матрицы $\mathbf{P}_{l \times l}^*$ отличаются от матриц $\mathbf{P}_{l-1 \times l}$ тем, что они содержат вероятности отнесения представленных на первом уровне i_1 -ых ТМП к i_l -ым ситуациям функционирования СТС на более высоком l -ом уровне иерархии через модели \mathbf{M}_{1,l,i_j,i_j} .

Начальные значения элементов матрицы вероятностей переходов $\mathbf{P}_{l-1 \times l}$ вида (2) определяются на основе экспертных знаний разработчиков СТС через модельные временные ряды ТМП в различных ситуациях функционирования СТС, а также на основе статистики отказов СТС по результатам сообщений о неисправностях и испытаний однотипных СТС [6, 32, 33]. В процессе обучения байесовской сети на реальных данных ТМИ о функционировании СТС матрицы $\mathbf{P}_{l-1 \times l}$ подлежат коррекции.

Вектора $\mathbf{P}_l^* = [P_{l,i_j}^*]$ характеризуют вероятности появления i_l -ых ситуаций функционирования СТС на каждом l -ом уровне иерархии. При этом справедливы следующие соотношения:

$$\begin{cases} \mathbf{P}_5^* = \mathbf{P}_4^* \mathbf{P}_{4 \times 5} \\ \mathbf{P}_4^* = \mathbf{P}_3^* \mathbf{P}_{3 \times 4} \\ \mathbf{P}_3^* = \mathbf{P}_2^* \mathbf{P}_{2 \times 3} \end{cases}, \begin{cases} \mathbf{P}_5^* = \mathbf{P}_1^* \mathbf{P}_{1 \times 5}^* \\ \mathbf{P}_4^* = \mathbf{P}_1^* \mathbf{P}_{1 \times 4}^* \\ \mathbf{P}_3^* = \mathbf{P}_1^* \mathbf{P}_{1 \times 3}^* \\ \mathbf{P}_2^* = \mathbf{P}_1^* \mathbf{P}_{1 \times 2}^* \end{cases}. \quad (3)$$

Соотношения (3) описывают различные варианты взаимосвязи ТМП и ситуаций функционирования СТС на различных уровнях ее иерархической структуры и являются основой агрегирования

разнородной информации, полученной по результатам обработки ТМП и знаний экспертов. Они предполагают равенство вероятностей достижения одних и тех же узлов дискретной байесовской сети прямым способом – с первого уровня на l -ый уровень – вследствие наличия ТМП, непосредственно связанных с соответствующими ситуациями функционирования СТС, и опосредованно – через узлы дискретной байесовской сети меньшего уровня иерархии – при наличии ТМП, характеризующих ситуации функционирования СТС косвенным образом.

С учетом обобщения соотношений (3) получим ограничения на значения вероятностей возникновения ситуаций функционирования на каждом уровне иерархии представления СТС в виде равенств:

$$\begin{cases} \mathbf{P}_{l_1}^* = \mathbf{P}_{l_1-1}^* \mathbf{P}_{l_1-1 \times l_1}, & 3 \leq l_1 \leq 5 \\ \mathbf{P}_{l_2}^* = \mathbf{P}_1^* \mathbf{P}_{1 \times l_2}^*, & 2 \leq l_2 \leq 5 \end{cases}, \quad (4)$$

где l_1 и l_2 – номера уровней иерархии.

В процессе обучения модели распознавания ситуаций функционирования СТС на основе дискретной байесовской сети необходимо обеспечить максимальное значение апостериорных вероятностей возникновения заданных j_l -ых ситуаций функционирования СТС на выходе каждого l -го уровня иерархии:

$$\begin{cases} P_{l,j_l}^*(\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*) = 1 - \sum_{\substack{i=1, \\ i \neq j_l}}^{N_l} P_{l,i}(\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*) \prod_{\substack{r=1, \\ r \neq i}}^{N_l} (1 - P_{l,r}(\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*)) \\ \{\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*\}_{opt} = \arg \max_{\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*} P_{l,j_l}^* \end{cases}, \quad (5)$$

где $\{\mathbf{P}_{l-1 \times l_1}^*, \mathbf{P}_{1 \times l_2}^*\}_{opt}$ – скорректированные матрицы вероятностей переходов, позволяющие получить максимальное значение апостериорной вероятности на выходе каждого l -го уровня агрегированной модели.

Для принятия решений при техническом диагностировании СТС внешней подсистемой технического диагностирования наиболее предпочтительным является критерий Вальда. Это обусловлено особенностью процесса отработки и анализа «тяжелых» отказов СТС, характеризующегося непрерывным «исследованием» неисправной СТС и получением внешней подсистемой технического

диагностирования дополнительной ТМИ для оценки ее технического состояния. В этой связи критерий Вальда можно считать критерием, наиболее адекватным существующим особенностям процесса диагностирования СТС. Это согласуется с оценками экспертов по техническому диагностированию СТС и находит свое выражение в широком использовании данного критерия при техническом диагностировании КА [34].

Для реализации принятия решений по критерию Вальда в агрегированной модели, представленной на рисунке 1, на выходе пятого уровня модели используется N_5 пороговых устройств, на которые поступают условные вероятности возникновения соответствующих РС из базы данных СТС. В каждое i_5 -ое пороговое устройство вводятся два порога a_{i_5} и b_{i_5} , которые обеспечивают заданный уровень ошибок первого и второго рода при распознавании соответствующей ситуации функционирования СТС. В зависимости от значения вероятностей $P_{i_5,k}$ на выходе пятого уровня по факту пересечения нижнего или верхнего порога, принимается решение об i_5 -ой ситуации функционирования СТС. После того, как по всем возможным i_5 -ым альтернативам получены значения условных вероятностей $P_{i_5,k}$, осуществляется принятие решения в пользу одной или нескольких наиболее вероятных согласно критерию максимума апостериорной вероятности ситуаций функционирования СТС.

Использование критерия Вальда расширяет перечень настраиваемых параметров пятого уровня агрегированной модели развития ситуаций функционирования СТС параметрами пороговых устройств $a_1, \dots, a_{N_5}, b_1, \dots, b_{N_5}$. При этом решение на выходе пятого уровня по всем альтернативам не будет сформировано до тех пор, пока не будет решения по каждой альтернативе на основе критерия Вальда. Поскольку матрицы (1) и (2) являются сильно разреженными и имеются ограничения на значения вероятностей каждого уровня иерархии в виде равенств (4), то существует нелинейная зависимость апостериорной вероятности принятия решения в пользу i_5 -ой ситуации функционирования от установленных пороговых значений. Становится возможной структурная оптимизация уровня подсистем агрегированной модели. При этом наиболее адекватными методами настройки элементов матриц (2) и коррекции элементов матриц (1) являются методы стохастической оптимизации [35,36].

Стохастическое обучение дискретной байесовской сети представляет собой процедуру «обучения с учителем» и включает два взаимосвязанных процесса. Это процесс оптимизации элементов

матриц (2) и коррекции матриц (1) по критерию максимума апостериорной вероятности на выходе каждого уровня СТС, а также процесс коррекции матриц (2) через обратную связь и матрицы (1), полученные в результате обработки ТМП. Результатом процесса обучения является агрегирование экспертной информации и результатов обработки ТМП в виде обновленных матриц (1) и (2), а также новые приобретенные возможности по распознаванию типовых ситуаций функционирования из базы данных СТС и локализации неисправных аппаратных модулей. Уточненные экспертные знания в виде матриц (2) могут быть использованы для решения задачи диагностирования СТС при отсутствии отдельных ТМП или при наличии значительных пропусков значений во временных рядах ТМП. Для обучения используются ситуации функционирования элементов СТС на каждом уровне сети и матрицы вторичных вероятностных признаков (1), полученные в ходе обработки ТМП, содержимое которых в процессе обучения лишь незначительно корректируется.

Решение задачи (5) осуществляется, исходя из целевой функции:

$$E_1(\mathbf{P}_{1 \times l_1}, \mathbf{P}_{1 \times l_2}^*, a_1 \dots a_{N_5}, b_1 \dots b_{N_5}) = \frac{\sum_{l=2}^5 \sum_{j_l=1}^{N_l} \left| P_{l,j_l}^* (\mathbf{P}_{1 \times l_1}) - P_{l,j_l}^* (\mathbf{P}_{1 \times l_2}^*) \right|}{\sum_{l=2}^5 \sum_{j_l=1}^{N_l} P_{l,j_l}^*}. \quad (6)$$

В выражении (6) для целевой функции E_1 числитель определяет сумму всех невязок между апостериорными значениями вероятностей на выходе каждого уровня иерархии, полученных на основе экспертных оценок вероятностей переходов, и оценок соответствующих вероятностей, полученных по результатам обработки ТМП. Знаменатель представляет собой сумму всех апостериорных вероятностей возникновения j_l -ых ситуаций функционирования СТС на выходе каждого l -го уровня иерархии дискретной байесовской сети, связанных с анализируемой рассмотренной ситуацией функционирования из базы данных СТС.

Целевая функция E_1 , определяемая выражением (6), позволяет выполнить условие максимизации апостериорной вероятности правильного распознавания ситуаций функционирования СТС при ограничениях в виде равенств (4) с учетом использования критерия Вальда:

$$\left(\mathbf{P}_{1-1 \times l_1}, \mathbf{P}_{1 \times l_2}^* \right)_{opt} = \arg \min_{\substack{\mathbf{P}_{1-1 \times l_1}, \mathbf{P}_{1 \times l_2}^* \\ a_1 \dots a_{N_5}, \\ b_1 \dots b_{N_5}}} E_1(\mathbf{P}_{1-1 \times l_1}, \mathbf{P}_{1 \times l_2}^*) \quad \left. \begin{array}{l} \mathbf{P}_{1 \times l_2}^* \in [P_{1 \times l_2, \min}, P_{1 \times l_2, \max}] \\ \mathbf{P}_{1-1 \times l_1} \in [P_{1-1 \times l_1, \min}, P_{1-1 \times l_1, \max}] \\ 3 \leq l_1 \leq 5, 2 \leq l_2 \leq 5, \end{array} \right\} \quad (7)$$

Задача (7) решается при заданных допусках $[P_{1 \times l_2, \min}, P_{1 \times l_2, \max}]$, $[P_{1-1 \times l_1, \min}, P_{1-1 \times l_1, \max}]$ на значения варьируемых вероятностей переходов дискретной байесовской сети, определяемых разработчиками по результатам моделирования, проведения различного рода испытаний и эксплуатации СТС. При этом интервал допустимых вторичных вероятностных признаков $[P_{1 \times l_2, \min}, P_{1 \times l_2, \max}]$ значительно меньше, чем интервал $[P_{1-1 \times l_1, \min}, P_{1-1 \times l_1, \max}]$.

Результатом решения задачи оптимизации (6) является уточненный набор вероятностей межуровневых переходов дискретной байесовской сети и переходов с первого уровня на остальные уровни. Для оценки качества решения задачи технического диагностирования, распознавания ситуаций функционирования СТС, используется показатель достоверности принятия решения в пользу j_5 -ой ситуации D_{j_5} , определяемой на основе полной группы возможных решений на выходе пятого уровня дискретной байесовской сети:

$$D_{j_5} = 1 - \sum_{\substack{\xi=1, \\ \xi \neq j_5}}^{N_5} P_{\xi} \prod_{\substack{r=1, \\ r \neq \xi}}^{N_5} (1 - P_r). \quad (8)$$

Значение показателя достоверности диагностирования СТС D_{j_5} для внешней системы контроля и диагностирования должно быть выше значения показателя достоверности диагностирования СТС встроенной системой контроля, которое задается в эксплуатационно-технической документации на СТС. Достаточность достижения требуемой достоверности диагностирования СТС можно определить через заданную надежность функционирования СТС, которая также определена нормативными документами [6].

Важным достоинством предлагаемого подхода является возможность обучения дискретной байесовской сети на типовых

значениях вероятностей переходов $\mathbf{P}_{l-1 \times l_1}$ и $\mathbf{P}_{1 \times l_2}^*$ из заданных допусков $\left[P_{l \times l_2, \min}, P_{l \times l_2, \max} \right]$ и $\left[P_{l_1-1 \times l_1, \min}, P_{l_1-1 \times l_1, \max} \right]$, определяемых экспертами, даже в отсутствии реальной ТМИ. Это важно на начальном этапе применения диагностической модели и эксплуатации СТС. Для практической реализации стохастического обучения рассмотренной дискретной байесовской сети в модели использован рандомизированный алгоритм стохастической аппроксимации, который применяется при решении задач многопараметрической оптимизации с большим количеством настраиваемых параметров и сильно разреженными матрицами настраиваемых параметров (матрицы $\mathbf{P}_{l_1-1 \times l_1}$ и $\mathbf{P}_{1 \times l_2}^*$) [35,36].

Полученные с помощью модели дискретной байесовской сети вероятности достижения вершин на каждом уровне иерархии \mathbf{P}_l^* можно рассматривать как обобщенные параметры состояния аппаратных модулей и подсистем СТС. Их можно использовать для решения задачи локализации проблемных функциональных узлов СТС, функционирование которых с наибольшей вероятностью отличается от штатного. С этой целью необходимо проанализировать вероятность достижения вершины графа состояний на каждом уровне иерархии и соотнести эту вероятность с вероятностью достижения вершины, характеризующей штатное функционирование данного элемента СТС. При этом на каждом уровне иерархии решается задача поиска множества $\{j_{l,nc}, P_{l,j_{l,nc}}^*\}$ номеров вершин, соответствующих нештатным ситуациям функционирования СТС, и вероятностей их достижения

$$P_{l,j_{l,nc}}^* = \arg \max_{P_{l,j_l}^*} E_{2,l,j_l} (P_{l,j_l}^*, P_{l,j_l,шф}^*), \forall j_l, \quad (9)$$

с максимизацией целевой функции

$$E_{2,l,j_l} = (P_{l,j_l}^* - P_{l,j_l,шф}^*) / P_{l,j_l,шф}^*. \quad (10)$$

В выражения (9) и (10) входят следующие вероятности: $P_{l,j_l,шф}^*$ – вероятность достижения j_l -ой вершины сети на уровне l , характеризующая ситуацию штатного функционирования СТС, P_{l,j_l}^* –

вероятность достижения j_l -ой вершины сети на уровне l , характеризующая текущую ситуацию функционирования СТС, $P_{l,j_l,nc}^*$ – вероятность достижения j_l -ой вершины сети на уровне l , характеризующая ситуацию нештатного функционирования СТС.

Для нормально функционирующих СТС решение задачи (9) и анализ изменения обобщенных параметров $P_{l,j_l,nc}^*$ во времени по результатам нескольких процедур технического диагностирования позволит выявить негативные тенденции в изменении технического состояния СТС и своевременно их парировать.

Представленное формализованное описание процесса технического диагностирования СТС по результатам обработки полных потоков ТМИ отражает направление совершенствования методов углубленного анализа технического состояния СТС, направленное на повышение достоверности их диагностирования и, тем самым, поддержание требуемой надежности функционирования. Не менее важным направлением обеспечения требуемой надежности функционирования СТС является автоматизация рассмотренного процесса технического диагностирования СТС и достижение требуемой оперативности углубленного анализа, определяемой периодичностью съема и обработки полных потоков ТМИ и временем принятия решения по управлению СТС.

4. Формализованное описание нелинейных динамических моделей изменения ТМП СТС с использованием искусственных нейронных сетей. Математический аппарат искусственных нейронных сетей позволяет реализовать нелинейные динамические модели изменения ТМП $\{M_{v,p}\}$ в широком диапазоне ситуаций функционирования аппаратных модулей и подсистем СТС. Их использование дает возможность повысить достоверность распознавания ситуаций функционирования СТС в условиях значительной неопределенности значений и существенно нелинейного характера изменения ТМП при возникновении нештатных ситуаций. При этом, как показывает теория и практика применения искусственных нейронных сетей для решения задач обработки временных рядов ТМП, достаточным является использование классических нейронных сетей с двумя скрытыми слоями [21, 24].

В простейшем случае модели изменения ТМП $\{M_{1,l,i_1,i_l}\}$ могут быть заданы набором допусков $[D_{n,i_1,i_l}; D_{v,i_1,i_l}]$, в пределах которых изменяются i_1 -ые ТМП в i_l -ых ситуациях функционирования аппаратных модулей и подсистем СТС. В предположении гауссовской

статистики значений ТМП, вероятность P_{1,l,i_1,i_l} отнесения i_1 -го ТМП к i_l -ой ситуации функционирования определяется выражением:

$$P_{1,l,i_1,i_l} = \Phi\left(\frac{D_{в,i_1,i_l} - \bar{x}_{i_1}}{\sigma_{i_1}}\right) - \Phi\left(\frac{D_{н,i_1,i_l} - \bar{x}_{i_1}}{\sigma_{i_1}}\right), \quad (11)$$

где \bar{x}_{i_1} – математическое ожидание i_1 -го анализируемого ТМП; σ_{i_1} – среднеквадратическое отклонение значений i_1 -го анализируемого ТМП от математического ожидания \bar{x}_{i_1} ; $[D_{н,i_1,i_l}; D_{в,i_1,i_l}]$ – доверительный интервал (допуск) на значения i_1 -го ТМП в i_l -ой ситуации функционирования аппаратных модулей и подсистем СТС, определяемый в соответствии со статистическими критериями с доверительной вероятностью $P_{\text{дов}}$ через модельные значения среднеквадратического отклонения σ_{i_1,i_l} и математического ожидания \bar{x}_{i_1,i_l} i_1 -го ТМП в i_l -ой ситуации функционирования; $D_{н,i_1,i_l}$, $D_{в,i_1,i_l}$ – нижняя и верхняя границы допуска на значения i_1 -го ТМП в i_l -ой ситуации функционирования соответственно;

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{x^2}{2}\right) dx$ – функция Лапласа. Величина $P_{\text{дов}}$

определяется исходя из требуемых достоверности контроля технического состояния и надежности функционирования СТС [37].

В качестве моделей $\{M_{1,l,i_1,i_l}\}$ также могут быть использованы нелинейные динамические модели изменения ТМП в проекциях их фазового пространства с реализацией парциальной нелинейной фильтрации [38]. В данной статье для разработки нелинейных моделей $\{M_{1,l,i_1,i_l}\}$ используются искусственные нейронные сети.

Структурная схема нейросетевой модели изменения ТМП в виде многослойной нейронной сети представлена на рисунке 2.

Входной вектор $X_{1,l,i_1,i_l,\varphi}$ нейронной сети содержит значения дискретных отсчетов i_1 -ых ТМП в i_l -ых ситуациях функционирования СТС в пределах φ -го «окна наблюдения» длины L дискретных отсчетов:

$$X_{1,l,i_1,i_l,\varphi} = [x_{1,l,i_1,i_l,k}, x_{1,l,i_1,i_l,k-1}, \dots, x_{1,l,i_1,i_l,k-L+1}].$$

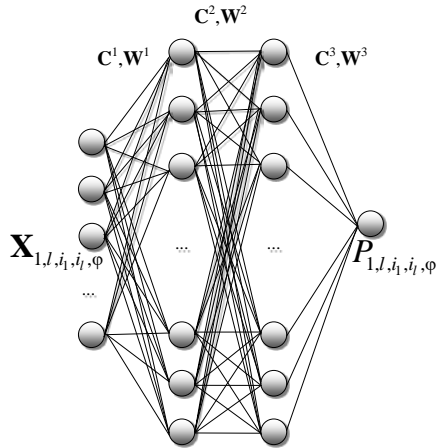


Рис. 2. Структурная схема нейросетевой модели изменения ТМП в заданной ситуации функционирования СТС

Каждый слой нейронной сети характеризуется матрицей весовых коэффициентов \mathbf{W}^μ и вектором смещений \mathbf{C}^μ , где μ – номер слоя, $\mu=1\dots3$. Выход сети имеет единичную размерность. Значение его, в ряде случаев, можно трактовать как условную вероятность отнесения текущего φ -го участка временной реализации i_l -го ТМП к i_l -ой ситуации функционирования аппаратных модулей и подсистем СТС $P_{1,l,i_l,i_l,\varphi}$. Функционирование представленной многослойной нейронной сети включает этапы обучения и применения.

На этапе обучения многослойных нейронных сетей минимизируется среднеквадратическая ошибка

$$E_{1,l,i_l,i_l} = \sum_{\varphi} \left(P_{1,l,i_l,i_l,\varphi} - P_{1,l,i_l,i_l,\text{зад}} \right)^2, \quad (12)$$

где $P_{1,l,i_l,i_l,\text{зад}}$, $P_{1,l,i_l,i_l,\varphi}$ – заданное и текущее значения вероятностей на выходе многослойной нейронной сети. Вероятности $P_{1,l,i_l,i_l,\text{зад}}$ принимают единичное значение «1» – при предъявлении сети на этапе обучения φ -ых участков множества реализаций ТМП, соответствующих i_l -ой ситуации функционирования аппаратного модуля или подсистемы СТС. Нулевое значение «0» – они имеют при предъявлении сети на этапе обучения φ -ых участков множества

реализаций ТМП всех остальных ситуаций функционирования данного аппаратного модуля или подсистемы СТС, а также множества произвольных случайных реализаций, формируемых в пределах телеметрической шкалы на этапе обучения многослойной сети. Для обучения модели используется алгоритм наискорейшего спуска или алгоритм Левенберга-Марквардта [24].

На этапе применения многослойных нейронных сетей определяется выходное значение вероятности $P_{1,l,i_1,i_l,\varphi}$:

$$P_{1,l,i_1,i_l,\varphi} = \mathbf{f}\left(\mathbf{f}\left(\mathbf{f}\left(\mathbf{X}_{i_1,i_l,\varphi} \mathbf{W}^1 + \mathbf{C}^1\right) \mathbf{W}^2 + \mathbf{C}^2\right) \mathbf{W}^3 + \mathbf{C}^3\right), \quad (13)$$

где $\mathbf{f}(\cdot)$ – нелинейный функционал преобразования вектора выходных значений каждого слоя, основанный на логистической функции активации. Для принятия решения о текущей ситуации функционирования аппаратного модуля или подсистемы СТС по i_1 -ым ТМП с использованием нейросетевых моделей $\{\mathbf{M}_{1,l,i_1,i_l}\}$ необходимо осуществить:

- перемещение «окна наблюдения» в пределах реализации каждого i_1 -го ТМП по временной шкале;
- распознавание анализируемого φ -го участка временных рядов ТМП;
- накопление полученных результатов распознавания в пределах всей анализируемой временной реализации ТМП;
- нормализацию полученной суммы и получение вероятностей отнесения всей реализации i_1 -го ТМП к каждой i_l -ой ситуации функционирования:

$$P_{1,l,i_1,i_l} = \frac{\sum_{\varphi} P_{1,l,i_1,i_l,\varphi}}{\sum_{i_l} \sum_{\varphi} P_{1,l,i_1,i_l,\varphi}}. \quad (14)$$

Полученные в результате моделирования вероятности P_{1,l,i_1,i_l} образуют матрицы вероятностей переходов многоуровневой иерархической модели на основе дискретной байесовской сети $\mathbf{P}_{1 \times l}^*$ (1).

5. Основные этапы реализации и результаты применения подхода к техническому диагностированию СТС с использованием моделей в виде гибридных сетевых структур. Практическая

реализация предлагаемого подхода к техническому диагностированию СТС с использованием моделей в виде гибридных сетевых структур включает два этапа: обучение диагностического комплекса и его применение по назначению для технического диагностирования СТС.

На этапе обучения диагностического комплекса выполняются следующие шаги:

1. Получение статистической базы динамических реализаций i_1 -ых ТМП в виде множества их значений $\{x_{1,i_1}\}$ в i_1 -ых ситуациях функционирования аппаратных модулей и подсистем СТС, а также реализаций ТМП, сформированных на основе знаний экспертов в области функционирования данной СТС. Разбиение экспертом всех данных полученной базы на три выборки: обучающую, проверочную и тестовую.

2. Формирование множества моделей изменения i_1 -ых ТМП в i_1 -ых ситуациях функционирования аппаратных модулей и подсистем СТС $\{M_{1,l,i_1,i_1}\}$ на основе полученных реализаций ТМП $\{x_{1,i_1}\}$ с использованием многослойных нейронных сетей и целевой функции (12).

3. Сопоставление реализаций ТМП $\{x_{1,i_1}\}$ из обучающей, проверочной и тестовой выборок с полученными моделями их изменения $\{M_{1,l,i_1,i_1}\}$ и формирование результатов данного сопоставления в виде множеств вторичных вероятностных признаков $\{P_{1,l,i_1,i_1}\}_o$, $\{P_{1,l,i_1,i_1}\}_n$ и $\{P_{1,l,i_1,i_1}\}_t$ согласно (13) и (14).

4. Формирование структуры многоуровневой иерархической модели M_2 на основе информации о взаимосвязях между ситуациями функционирования СТС на различных иерархических уровнях ее представления в виде дискретной байесовской сети.

5. Формирование множества вероятностей переходов $\{P_{-1,l,i_1-1,i_1}\}$ между различными ситуациями функционирования СТС на всех иерархических уровнях ее представления, определение допустимых значений на изменение вероятностей переходов $\left[P_{1-1 \times l_1, \min} \cdot P_{1-1 \times l_1, \max} \right]$ и $\left[P_{1 \times l_2, \min} \cdot P_{1 \times l_2, \max} \right]$, $3 \leq l_1 \leq 5$, $2 \leq l_2 \leq 5$, с использованием знаний экспертов в области функционирования СТС.

6. Ввод в дискретную байесовскую сеть полученных вторичных вероятностных признаков $\{P_{1,l,i_1,i_1}\}_o$, соответствующих обучающей выборке, а также начальных значений вероятностей $\{P_{l-1,l,i_1-1,i_1}\}$ и

параметров пороговых устройств $a_1, \dots, a_{N_5}, b_1, \dots, b_{N_5}$.

7. Стохастическое обучение дискретной байесовской сети с использованием алгоритма стохастической аппроксимации в соответствии с целевой функцией E_1 (6) и (7) и ограничениями в виде равенств (3) и (4). Получение в результате обучения множества вероятностей $\{P_{l-1,l,i_{l-1},i_l}\}^*$ и параметров $a^*_1, \dots, a^*_{N_5}, b^*_1, \dots, b^*_{N_5}$.

8. Проверка результата обучения дискретной байесовской сети на множестве вторичных вероятностных признаков $\{P_{1,l,i_l,i_l}\}_n$ проверочной выборки и множестве полученных в результате обучения вероятностей $\{P_{l-1,l,i_{l-1},i_l}\}^*$, вычисление значения целевой функции E_1 вида (6). Если значение целевой функции E_1 уменьшается в сравнении со значением, полученным на предыдущем цикле обучения, то фиксируются текущие параметры дискретной байесовской сети. Если количество циклов обучения не превысило заданного значения $N_{\text{о зад}}$, то осуществляется выполнение шага 7, иначе – шага 9.

9. Проверка результата обучения дискретной байесовской сети на тестовой выборке $\{P_{1,l,i_l,i_l}\}_t$ с использованием полученного в ходе обучения множества вторичных вероятностных признаков $\{P_{l-1,l,i_{l-1},i_l}\}^*$: вычисление выходных вероятностей на каждом слое дискретной байесовской сети $\{P_{l,i}\}$, оценка достоверности D_{j_5} принятия решения в пользу j_5 -ой ситуации функционирования СТС в соответствии с выражением (8). Если значение D_{j_5} ниже заданного значения $D_{\text{зад}}$, то осуществляется переход к выполнению шага 7, иначе – завершение этапа обучения дискретной байесовской сети.

На этапе применения диагностического комплекса выполняются следующие шаги:

1. Получение новых реализаций i_1 -ых ТМП в виде множества отсчетов временных рядов $\{x_{1,i_l}\}$.

2. Сопоставление полученных реализаций i_1 -ых ТМП с моделями изменения ТМП $\{\mathbf{M}_{1,l,i_l,i_l}\}$ в i_l -ых ситуациях функционирования аппаратных модулей и подсистем СТС и формирование множества текущих вероятностных оценок данного сопоставления $\{P_{1,l,i_l,i_l}\}$.

3. Ввод в дискретную байесовскую сеть вероятностей $\{P_{1,l,i_l,i_l}\}$, а также вероятностей $\{P_{l-1,l,i_{l-1},i_l}\}^*$ и параметров пороговых устройств

$a^*_{1}, \dots, a^*_{N_5}, b^*_{1}, \dots, b^*_{N_5}$, полученных на этапе обучения.

4. Вывод информации о вероятностях ситуаций функционирования СТС на выходе дискретной байесовской сети P^*_{5,j_5} .

Определение наиболее вероятной ситуации функционирования СТС.

5. Определение значений обобщенных параметров $P^*_{l,j_l,nc}$ элементов СТС и локализация наиболее проблемных из них, удовлетворяющих критерию (9).

6. Выявление тенденций развития негативных процессов изменения технического состояния СТС на основе анализа значений обобщенных параметров $P^*_{l,j_l,nc}$, полученных по результатам обработки полных потоков ТМИ при проведении последующих сеансов диагностирования и управления ее техническим состоянием.

В результате технического диагностирования СТС с использованием рассмотренной диагностической модели в виде гибридных сетевых структур возможны следующие исходы:

1. В одном или нескольких выходных узлах дискретной байесовской сети принято решение о наличии рассмотренной ситуации функционирования из базы данных СТС. Формируется объяснение результата распознавания ситуации функционирования в виде последовательности вершин графической модели $\{P^*_{l,j_l,nc}\}$ на каждом уровне иерархического представления СТС и локализуется неисправный элемент СТС.

2. В выходных узлах дискретной байесовской сети не принято решение о рассмотренной ситуации функционирования из базы данных СТС. Осуществляется ранжирование вершин $\{P^*_{l,j_l,nc}\}$ в соответствии с критерием (9) на каждом уровне иерархического представления СТС и локализуются наиболее проблемные элементы СТС.

3. На основе анализа выявленных проблемных элементов СТС и обобщенных параметров $\{P^*_{l,j_l,nc}\}$ по результатам последующих сеансов диагностирования и управления техническим состоянием СТС формируется вывод о развитии негативных тенденций в изменении технического состояния и функционировании СТС.

В первом случае в результате технического диагностирования СТС осуществляется поддержка принятия решения во внешней системе управления СТС при реализации ситуационного управления ее техническим состоянием и повышается качество управления СТС с

точки зрения оперативности принятия решения. Во втором и третьем случае в результате технического диагностирования СТС выявляются возможные негативные тенденции в изменении технического состояния и функционирования СТС и проблемные элементы СТС, что позволяет реализовать перспективные способы проактивного управления СТС [8–10, 39].

Следует отметить, что при реализации предлагаемого подхода важная роль отводится экспертам в области функционирования СТС, разрабатывающим структуру дискретной байесовской сети и определяющим диапазоны изменения вероятностей переходов между ситуациями функционирования аппаратных модулей и подсистем в СТС. При этом существует три способа получения экспертных знаний о вероятностях переходов $\{P_{l-1,i,l-1,i}\}$. Первый способ основан на приблизительной оценке допустимых значений $[P_{l-1 \times l_1, \min} \cdot P_{l-1 \times l_1, \max}]$, исходя из анализа сообщений о неисправностях и результатов испытаний однотипных СТС. Второй способ основан на косвенной приблизительной оценке допустимых значений $[P_{l-1 \times l_1, \min} \cdot P_{l-1 \times l_1, \max}]$ в ходе предварительного имитационного эксперимента и стохастического обучения сети, наполненной вероятностями переходов $[P_{l \times l_2, \min} \cdot P_{l \times l_2, \max}]$, полученных на основе модельных ТМП, заданных экспертами для различных ситуаций функционирования СТС. Третий способ основан на комбинировании обоих перечисленных способов.

6. Результаты моделирования процесса технического диагностирования СТС на примере бортовой аппаратуры КА. Проведено моделирование процесса технического диагностирования СТС с использованием предлагаемого подхода. В качестве объекта технического диагностирования рассматривалась бортовая система приема и передачи информации. В иерархическом представлении КА, как СТС, данная система относится к уровню функциональных систем КА. Поэтому модель технического диагностирования данной системы, в отличие от обобщенной модели, приведенной на рисунке 1, имеет более простую четырехуровневую структуру: ТМП, аппаратные модули, подсистемы, функциональная система. Пороговые устройства используются на выходе четвертого уровня модели, на которые поступают условные вероятности возникновения текущей ситуации функционирования системы приема и передачи информации. Моделирование работы дискретной байесовской сети осуществлялось

в среде имитационного моделирования AnyLogic [40].

На рисунке 3 представлено отображение диагностической модели системы приема и передачи информации в виде дискретной байесовской сети в среде AnyLogic. В модели учитываются 13 ТМП (вершины Q1_1 – Q1_13), полученные в ходе сеансов управления с КА. Временные реализации их преобразуются во вторичные вероятностные признаки ситуаций функционирования анализируемой системы, и 23 ситуации функционирования на всех уровнях формализованного описания функционирования системы в виде дискретной байесовской сети (вершины Q2_1–Q4_2). Выход дискретной байесовской сети представлен двумя вершинами, характеризующими штатное функционирование системы (вершина Q4_1) и наличие нештатной ситуации (вершина Q4_2). Диапазоны изменения вероятностей переходов $\{P_{l-1,l,i_{l-1},i_l}\}$ и их начальные значения определены приближенно на основе логики функционирования аппаратуры. Для формирования вероятностей $\{P_{l,l,i_l,i_l}\}$ отнесения ситуации функционирования системы приема и передачи информации КА на каждом уровне иерархии к типовым ситуациям функционирования по реализациям ТМП использованы динамические модели изменения ТМП на основе многослойных нейронных сетей. Размер «окна наблюдения» каждого ТМП варьировался от $L=50$ до $L=150$ отсчетов. Нейронные сети имели количество входных нейронов – от 50 до 100, количество скрытых слоев – 2, количество нейронов в каждом скрытом слое – 10, количество выходных нейронов – 1. При обучении нейронных сетей использованы критерий минимума среднеквадратической ошибки и алгоритм наискорейшего спуска [24].

Полученные значения вероятностей переходов $\{P_{l,l,i_l,i_l}\}$ и $\{P_{l-1,l,i_{l-1},i_l}\}$ загружались в дискретную байесовскую сеть, представленную на рисунке 3.

Стохастическое обучение сети реализовано на основе имитационного эксперимента, в ходе которого на вход модели дискретной байесовской сети в среде AnyLogic одновременно поступали заявки, распространяющиеся вдоль ее графа в соответствии с заданными вероятностями переходов. При этом происходила фиксация количества заявок, достигших каждой вершины на каждом уровне иерархии сети, и оценивались вероятности их достижения $P_{l,l}^*$. В ходе обучения модели происходило изменение вероятностей переходов в матрицах $\mathbf{P}_{l-1 \times l}$ и осуществлялась коррекция вероятностей переходов в матрице $\mathbf{P}_{l \times l}$ с целью оптимизации целевой функции (6)

при ограничениях в виде равенств (4) на основе алгоритма стохастической аппроксимации.

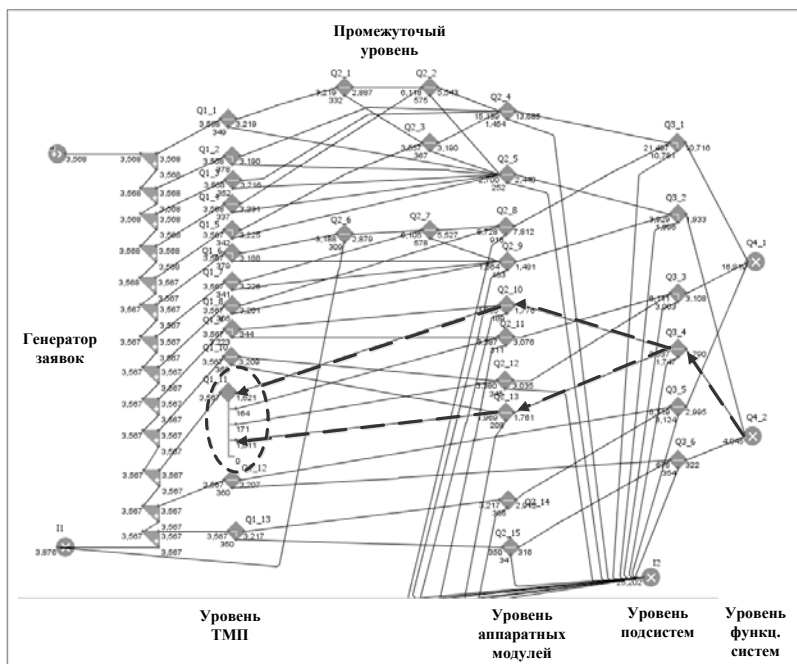


Рис. 3. Отображение в среде AnyLogic диагностической модели системы приема и передачи информации и результата локализации ее проблемного аппаратного модуля – источника вторичного электропитания приемника

Обученная модель позволила распознать рассмотренные ситуации штатного функционирования и нештатные ситуации системы приема и передачи информации на каждом уровне ее иерархии и решить задачу (9) локализации ее проблемных элементов. Кроме того, модель позволила выявить нерассмотренную ситуацию функционирования анализируемой системы и локализовать проблемный аппаратный модуль – источник вторичного электропитания приемника, характеризуемый нестабильностью напряжения питания и изменением его значения вблизи нижней границы допуска. На рисунке 3 пунктирной линией показан результат локализации данного аппаратного модуля, связанного со следующими вершинами диагностической сети: Q4_2, Q3_4, Q2_10, Q2_13, Q1_11.

На рисунке 4 представлены результаты распознавания ситуаций

функционирования одного из элементов приемника системы приема и передачи информации – аппаратного модуля шумовой автоматической регулировки усиления (ШАРУ) по одному из ее ТМП – «Напряжение регулировки». На рисунке 4а представлены значения вероятностей $P_{1,l,i\hat{i}}$ отнесения реализации анализируемого ТМП «Напряжение регулировки» к ситуациям функционирования «Штатное функционирование» (номер ситуации – 1) и «Обрыв в цепи обратной связи» (номер ситуации – 2), полученные с использованием метода допусков (столбцы 1) и на основе нейросетевой динамической модели изменения ТМП (столбцы 2) в отсутствии возмущений и шумов.

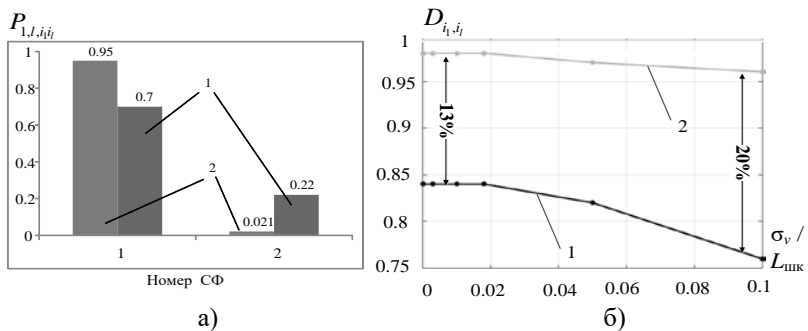


Рис. 4. Результаты распознавания ситуаций функционирования аппаратного модуля ШАРУ приемника системы приема и передачи информации по одному из ее ТМП – «Напряжение регулировки»: а) значения вероятностей отнесения реализации анализируемого ТМП «Напряжение регулировки» к ситуациям функционирования «Штатное функционирование» (номер ситуации – 1) и «Обрыв в цепи обратной связи» (номер ситуации – 2) б) зависимости показателя достоверности принятия решения о возможных ситуациях

На рисунке 4б представлены зависимости показателя достоверности $D_{i,i\hat{i}}$ принятия решения о возможных i -ых ситуациях функционирования для рассмотренного аппаратного модуля ШАРУ.

Приведенные зависимости показателя достоверности $D_{i,i\hat{i}}$ от величины $\sigma_v / L_{шк}$ получены по результатам формирования вторичных вероятностных признаков для каждой ситуации функционирования аппаратного модуля ШАРУ на основе определения принадлежности ТМП «Напряжение регулировки» допускам (кривая 1) и на основе нейросетевых динамических моделей изменения ТМП (кривая 2).

На рисунке 5 представлены результаты распознавания ситуации

функционирования системы приема и передачи информации по совокупности ее 13 ТМП. На рисунке 5а представлена экспериментальная зависимость изменения нормированных значений целевой функции (6) от дискретного времени k обучения байесовской сети. На рисунке 5б представлены зависимости показателя достоверности D_{j_4} распознавания анализируемой ситуации функционирования системы приема и передачи информации по совокупности ТМП с использованием дискретной байесовской сети в условиях динамических возмущений во временных рядах ТМП от параметра среднеквадратического отклонения гауссовского возмущения σ_v , нормированного к шкале ТМП $L_{\text{шк}}$.

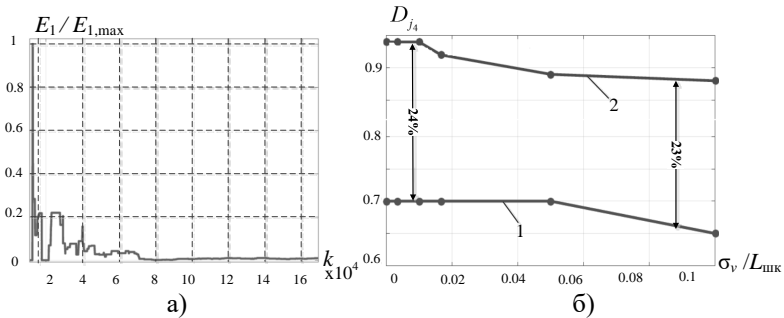


Рис. 5. Результаты распознавания ситуации функционирования системы приема и передачи информации по совокупности ее ТМП:
 а) экспериментальная зависимость изменения нормированных значений целевой функции (6) от дискретного времени k обучения байесовской сети;
 б) зависимости показателя достоверности распознавания анализируемой ситуации

В результате моделирования показана возможность получения выигрыша до 20% в достоверности распознавания ситуаций функционирования анализируемой функциональной системы КА при использовании в диагностической сети нейросетевых моделей изменения ТМП в сравнении с моделями на основе допусков. Нейросетевые модели, реализующие квазиоптимальную фильтрацию отдельных ТМП, позволили повысить достоверность распознавания ситуаций функционирования аппаратных модулей бортовой аппаратуры при увеличении интенсивности помех, действующих в радиоканале передачи ТМИ с КА, а также при наличии существенных возмущений и нелинейного характера временных рядов ТМП, характеризующих некоторые нештатные ситуации.

7. Заключение. Предложенный подход к техническому диагностированию СТС соответствует многомодельному описанию процессов их функционирования на различных иерархических уровнях представления и гибридному построению интеллектуальных систем обработки информации, предполагающему совместное использование подсистем статистического распознавания и логико-вероятностного вывода. Практическая реализация рассмотренных моделей в диагностических комплексах СТС различного назначения позволяет эффективно решать комплекс взаимосвязанных задач их технического диагностирования:

- обработки полных потоков ТМИ и отнесения текущей ситуации функционирования к множеству известных штатных и нештатных ситуаций из базы данных СТС;

- обнаружения ситуаций функционирования, не рассмотренных в базе данных СТС;

- уточнения штатной ситуации функционирования при ошибочном диагностировании СТС встроенными средствами технического диагностирования;

- локализации неисправных и проблемных элементов с выдачей во внешние системы контроля и технического диагностирования и управления СТС предложений по их детальному исследованию;

- выявления негативных тенденций в изменении технического состояния СТС по результатам локализации наиболее проблемных элементов.

Дальнейшее развитие рассмотренного подхода к техническому диагностированию СТС связано с поиском адекватного математического аппарата для разработки динамических моделей изменения ТМП, позволяющих достичь более выраженных максимумов гистограммы распределения вероятностей отнесения реализаций ТМП к возможным ситуациям функционирования отдельных аппаратных модулей и подсистем СТС.

Литература

1. Поспелов Д.А. Ситуационное управление: теория и практика // М.: Едиториал УРСС. 2021. 288 с.
2. Мостовой Я.А. Управление сложными техническими системами: конструирование программного обеспечения спутников ДЗЗ // М.: Техносфера. 2016. 352 с.
3. Александровская Л. Н. Современные методы обеспечения безопасности сложных технических систем: учебник // М.: Логос. 2001. 208 с.
4. Кирилин А.Н., Ахметов Р.Н., Сологуб А.Н., Макаров В.П. Методы обеспечения живучести низкоорбитальных автоматических космических аппаратов

- зондирования Земли: математические модели, компьютерные технологии // М.: Машиностроение. 2010. 384 с.
5. Carlton A., Morgan R., Lohmeyer W., Cahoy K. Telemetry Fault-Detection Algorithms: Applications for Spacecraft Monitoring and Space Environment Sensing // *Journal of Aerospace Information Systems*. 2018. vol. 15. no. 5. pp. 239–252.
 6. Мальцев Г.Н., Назаров А.В., Якимов В.Л. Исследование процесса диагностирования бортовой аппаратуры автоматических космических аппаратов с использованием дискретно-событийной имитационной модели // *Труды СПИИРАН*. 2018. Вып. 1(56). С. 95–121.
 7. Hassanien A.E., Darwish A., Abdelghafar S. Machine learning in Telemetry Data Mining of Space Mission: Basics, Challenging and Future Directions // *Artificial Intelligence Review*. 2020. vol. 53. no. 5. pp. 3201–3230.
 8. Bakhmut A.D., Krylov A.V., Okhtilev M.Y. Proactive Management of Complex Objects Using Precedent Methodology // *Advances in Intelligent Systems and Computing*. 2019. vol. 764. pp. 298–307.
 9. Bakhmut A.D., Ustinov A.V., Koromyshlichenko V.N. and oth. Methods of Conceptual Modeling of Intelligent Decision Support Systems for Managing Complex Objects at All Stages of its Life Cycle // *Advances in Intelligent Systems and Computing*. 2019. vol. 875. pp. 171–180.
 10. Охтилев М.Ю., Мустафин Н.Г., Миллер В.Е., Соколов Б.В. Концепция проактивного управления сложными объектами: теоретические и технологические основы // *Известия высших учебных заведений. Приборостроение*. 2014. Т. 57. № 11. С. 7–15.
 11. Titkanloo H.N., Keramati A., Fekri R. Data Aggregation in Multi-Source Assessment Model Based on Evidence Theory // *Applied Soft Computing*. 2018. vol. 69. pp. 443–452.
 12. Luger G. *Artificial Intelligence: Structure and Strategies for Complex Problem Solving*. 5th ed. // Harlow: Addison Wesley. 2005. 903 p.
 13. Burnham K.P. *Model Selection and Multimodel Inference: a Practical Informationtheoretic Approach* // New York: Springer. 2002. 488 p.
 14. Микони С.В., Соколов Б.В., Юсупов Р.М. Квалиметрия моделей и полимодельных комплексов // М.: РАН. 2018. 314 с.
 15. Лоскутов А.И., Бянкин А.А., Семенюк Д.Б., Клыков В.А. Методика синтеза математической модели функционирования бортовой радиоэлектронной аппаратуры объектов ракетно-космической техники на основе полимодельного подхода // *Приборы и системы. Управление, контроль, диагностика*. 2019. № 9. С. 27–36.
 16. Шмелёв В.В., Самойлов Е.Б., Саноцкая Н.А. Полимодельный комплекс интеллектуального мониторинга состояния технических систем с компенсацией возмущающих факторов // *T-Comm: Телекоммуникации и транспорт*. 2018. Т. 12. №8. С. 44–51.
 17. Jimenez J., Schwartz S., Vingerhoeds R., Grabot B. Towards Multi-Model Approaches to Predictive Maintenance: A Systematic Literature Survey on Diagnostics and Prognostics // *Journal of Manufacturing Systems*. 2020. vol. 56. pp. 539–557.
 18. Zhang P., Li T., Wang G., Luo C. Multi-Source Information Fusion Based on Rough Set Theory: A Review // *Information Fusion*. 2021. vol. 68. pp. 85–117.
 19. Li Z., Tian L., Jiang Q., Yan X. Fault Diagnostic Method Based on Deep Learning and Multimodel Feature Fusion for Complex Industrial Processes // *Industrial & Engineering Chemistry Research*. 2020. vol. 59. pp. 18061–18069.

20. Shahraki A.F., Yadav O.P., Liao H. A Review on Degradation Modelling and Its Engineering Applications // *International Journal of Performability Engineering*. 2017. vol. 13. pp. 299–314.
21. Якимов В.Л. Прогнозирование параметров технического состояния стартового комплекса с использованием нейронных сетей // *Известия высших учебных заведений. Приборостроение*. 2006. № 7. С. 7–10.
22. Marushko E.E., Doudkin A.A. Ensembles of Neural Networks for Forecasting of Time Series of Spacecraft Telemetry // *Optical Memory and Neural Networks*. 2017. vol. 26. no. 1. pp. 47–54.
23. Емельянова Ю.Г., Фраленко В.П. Методы когнитивно-графического представления информации для эффективного мониторинга сложных технических систем // *Программные системы: теория и приложения*. 2018. Т. 9. № 4(39). С. 117–158.
24. Мальцев Г.Н., Назаров А.В., Якимов В.Л. Алгоритм максимизации энтропии обучающей выборки и его использование при разработке прогнозной модели дискретных состояний нелинейной динамической системы // *Информационно-управляющие системы*. 2015. № 2. С. 57–66.
25. Preece A. Asking «Why» in AI: Explainability of Intelligent Systems – Perspectives and Challenges // *Intelligent Systems in Accounting, Finance and Management*. 2018. vol. 25. pp. 63–72.
26. Marcot B.G., Penman T.D. Advances in Bayesian Network Modelling: Integration of Modelling Technologies // *Environmental Modelling and Software*. 2019. vol. 111. pp. 386–393.
27. Chen J., Pi D., Wu Z., Zhao X. Imbalanced Satellite Telemetry Data Anomaly Detection Model Based on Bayesian LSTM // *Acta Astronautica*. 2021. vol. 180. pp. 232–242.
28. Масленников Е.Д., Сулимов В.Б. Предсказания на основе байесовских сетей доверия: алгоритм и программная реализация // *Вычислительные методы и программирование*. 2010. Т. 11. № 4. С. 94–107.
29. Мусина В.Ф. Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков // *Труды СПИИРАН*. 2013. Вып. 1(24). С. 135–151.
30. Gao X., Guo Z., Ren H., Yang Y. Learning Bayesian Network Parameters Via Minimax Algorithm // *International Journal of Approximate Reasoning*. 2019. vol. 108. pp. 62–75.
31. Торопова А.В. Подходы к диагностике согласованности данных в байесовских сетях доверия // *Труды СПИИРАН*. 2015. Вып. 6(43). С. 156–178.
32. Смирнов В.А. Поиск неисправностей в бортовых системах управления в процессе приемочного контроля // *Информационно-управляющие системы*. 2013. № 2. С. 24–28.
33. Ларин В.П., Шелест Д.К. Формирование информационного обеспечения надежности бортовой аппаратуры на стадии проектирования // *Информационно-управляющие системы*. 2012. № 4. С. 93–97.
34. Ахметов Р.Н. Методы и модели автономного управления живучестью автоматических космических аппаратов дистанционного зондирования Земли // *Вестник Самарского государственного аэрокосмического университета*. 2008. № 2. С. 194–210.
35. Kresoja M., Luzanin Z., Stojkovska I. Adaptive Stochastic Approximation Algorithm // *Numerical Algorithms*. 2017. vol. 76. pp. 917–937.
36. Граничин О.Н., Измакова О.А. Рандомизированный алгоритм стохастической аппроксимации в задаче самообучения // *Автоматика и телемеханика*. 2005. № 8. С. 52–63.

37. Wangab D., Tiana L. Parametric Methods for Confidence Interval Estimation of Overlap Coefficients // *Computational Statistics & Data Analysis*. 2017. vol. 106. pp. 12–26.
38. Мальцев Г.Н., Якимов В.Л., Соловьев С.В., Лебедева Н.В. Первичная обработка телеметрической информации с использованием динамических моделей изменения параметров и парциальной нелинейной фильтрации // *Информационно-управляющие системы*. 2018. № 5. С. 22–34.
39. Andoga R., Fozo L., Judicak J., Breda R. Intelligent Situational Control of Small Turbojet Engines // *International Journal of Aerospace Engineering*. 2018. pp. 1–16.
40. Gorod A., Ireland V., Gunawan I., Hallo L. Evolving Toolbox for Complex Project Management // *Voca Raton: Auerbach Publications*. 2019. 570 p.

Якимов Виктор Леонидович — д-р техн. наук, доцент, доцент, кафедра телеметрических систем, комплексной обработки и защиты информации, Военно-космическая академия имени А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: моделирование сложных систем, нейросетевые технологии, техническая диагностика. Число научных публикаций — 55. yakim78@yandex.ru; Ждановская улица, 13, 197198, Санкт-Петербург, Россия; р.т.: +7(812)347-95-32.

Мальцев Георгий Николаевич — д-р техн. наук, профессор, профессор, кафедра космических радиотехнических систем, Военно-космическая академия имени А.Ф. Можайского (ВКА им. А.Ф. Можайского). Область научных интересов: обработка сигналов в радиотехнических и оптико-электронных информационных системах, космические радиотехнические комплексы управления, сбора и передачи информации. Число научных публикаций — 200. georgy_maltsev@mail.ru; Ждановская улица, 13, 197198, Санкт-Петербург, Россия; р.т.: +7(812)237-12-49.

V. YAKIMOV, G. MALTSEV
**HYBRID NETWORK STRUCTURES AND THEIR USE IN
DIAGNOSING COMPLEX TECHNICAL SYSTEMS**

Yakimov V., Maltsev G. Hybrid Network Structures and Their Use in Diagnosing Complex Technical Systems.

Abstract. An approach to the technical diagnostics of complex technical systems based on the results of telemetry information processing by an external monitoring and diagnostics system using hybrid network structures is proposed. The principle of constructing diagnostic complexes of complex technical systems is considered, which ensures the automation of the technical diagnostics process and is based on the use of models in the form of hybrid network structures for processing telemetric information, including multilayer neural networks and discrete Bayesian networks with stochastic learning. A model of changes in the parameters of complex technical systems technical state based on multilayer neural networks has been developed, which makes it possible to form a probabilistic assessment of attributing the current situation of complex technical system functioning to the set of functions considered situations according to individual telemetry parameters, and multilevel hierarchical model of complex technical systems technical diagnostics based on a discrete Bayesian network with stochastic learning, which allows aggregating the information received from neural network models and recognizing the current situation of complex technical system functioning. In the conditions of functioning emergencies of the complex technical system, according to the results of processing telemetric information, faulty functional units are localized and an explanation of the cause of the emergency is formed. The stages of complex technical systems technical diagnostics implementation using the proposed hybrid network structures in the processing of telemetric information are detailed. An example of using the developed approach to solving problems of spacecraft onboard system technical diagnostics is presented. The advantages of the proposed approach to the technical diagnostics of complex technical systems in comparison with the traditional approach based on analysis of telemetry parameters values belonging to the given tolerances are shown.

Keywords: complex technical system, technical diagnostics, reliability of functioning, bayesian network, stochastic learning.

Yakimov Victor — Ph.D., Dr.Sci., Associate Professor, Associate professor, Department of the space radio engineering systems, A.F. Mozhaysky military space academy. Research interests: modeling of complex systems, neural network technologies, technical diagnostics. The number of publications — 55. yakim78@yandex.ru; 13, Zhdanovskaya St., 197198, St. Petersburg, Russia; office phone: +7(812)347-95-32.

Maltsev Georgiy — Ph.D., Dr.Sci., Professor, Department of the space radio engineering systems, A.F. Mozhaysky military space academy. Research interests: processing of signals in radio engineering and optico-electronic information systems, space radio-technical complexes for control, collection and transmission of information. The number of publications — 200. georgy_maltsev@mail.ru; 13, Zhdanovskaya St., 197198, St. Petersburg, Russia; office phone: +7(812)237-12-49.

References

1. Pospelov D.A. *Situatsionnoe upravlenie: teoriia i praktika* [The Situational Management: Theory and Practice]. Moscow: Editorial URSS Publ., 2021. 288 p. (In Russ.).
2. Mostovoi Ia.A. *Upravlenie slozhnymi tekhnicheskimi sistemami: konstruirovaniye programmogo obespecheniya sputnikov DZZ* [Management of Complex Technical Systems: Software Design for Remote Sensing Satellites]. Moscow: Tehnosfera Publ., 2016. 352 p. (In Russ.).
3. Aleksandrovskaia L.N. *Sovremennyye metody obespecheniya bezopasnosti slozhnykh tekhnicheskikh sistem: uchebnyk* [Modern Methods of Ensuring the Safety of Complex Technical Systems: Textbook]. Moscow: Logos Publ., 2001. 208 p. (In Russ.).
4. Kirilin A.N., Akhmetov R.N., Sologub A.N., Makarov V.P. *Metody obespecheniya zhivuchesti nizkoorbital'nykh avtomaticheskikh kosmicheskikh apparatov zondirovaniya Zemli: matematicheskie modeli, komp'yuternyye tekhnologii* [Methods for Ensuring the Survivability of Low-Orbit Automatic Earth Sensing Spacecraft: Mathematical Models, Computer Technologies]. Moscow: Mashinostroenie Publ., 2010. 384 p. (In Russ.).
5. Carlton A., Morgan R., Lohmeyer W., Cahoy K. *Telemetry Fault-Detection Algorithms: Applications for Spacecraft Monitoring and Space Environment Sensing*. Journal of Aerospace Information Systems. 2018. vol. 15. no. 5. pp. 239–252.
6. Mal'tsev G.N., Nazarov A.V., Yakimov V.L. [The Study of the Diagnosing of Automatic Spacecraft On-board Equipment Using a Discrete-Event Simulation Model]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2018. Issue 1(56). pp. 95–121. (In Russ.).
7. Hassanien, A.E., Darwish A., Abdelghafar S. *Machine learning in Telemetry Data Mining of Space Mission: Basics, Challenging and Future Directions*. Artificial Intelligence Review. 2020. vol. 53. no. 5. pp. 3201–3230.
8. Bakhmut A.D., Krylov A.V., Okhtilev M.Y. *Proactive Management of Complex Objects Using Precedent Methodology*. Advances in Intelligent Systems and Computing. 2019. vol. 764. pp. 298–307.
9. Bakhmut A.D., Ustinov A.V., Koromyslichenko V.N. and oth. *Methods of Conceptual Modeling of Intelligent Decision Support Systems for Managing Complex Objects at All Stages of its Life Cycle*. Advances in Intelligent Systems and Computing. 2019. vol. 875. pp. 171–180.
10. Okhtilev M.Ju., Mustafin N.G., Miller V.E., Sokolov B.V. [The Concept of Management Complex Objects Proactive: Theoretical and Technological Foundations]. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie – Proceedings of higher educational institutions. Instrumentation*. 2014. vol. 57. no. 11. pp. 7–15. (In Russ.).
11. Titkanloo H.N., Keramati A., Fekri R. *Data Aggregation in Multi-Source Assessment Model Based on Evidence Theory*. Applied Soft Computing. 2018. vol. 69. pp. 443–452.
12. Luger G. *Artificial Intelligence: Structure and Strategies for Complex Problem Solving*. 5th ed. Harlow: Addison Wesley, 2005. 903 p.
13. Burnham K.P. *Model Selection and Multimodel Inference: a Practical Informationtheoretic Approach*. NewYork: Springer, 2002. 488 p.
14. Mikoni S.V., Sokolov B.V., Iusupov R.M. *Kvalimetriia modelei i polimodel'nykh kompleksov* [Qualimetry of Models and Polymodel Complexes]. Moscow: RAN Publ., 2018. 314 p. (In Russ.).
15. Loskutov A.I., Biankin A.A., Semeniuk D.B., Klykov V.A. [Methods for Onboard Radio electronic Equipment Functioning Mathematical Model Synthesis of Rocket Objects and Space Technology Based on Polymodel Approach]. *Pribery i sistemy*.

- Upravlenie, kontrol', diagnostika – Devices and Systems. Management, Control, Diagnostics. 2019. no. 9. pp. 27–36. (In Russ.).
16. Shmelev V.V., Samoilov E.B., Sanotskaia N.A. [Polymodel Complex for Intelligent Monitoring of Technical Systems State with Disturbing factors Compensation]. T-Comm: Telekommunikatsii i transport. – T-Comm: Telecommunications and Transport. 2018. vol. 12. no. 8. pp. 44–51. (In Russ.).
 17. Jimenez J., Schwartz S., Vingerhoeds R., Grabot B. Towards Multi-Model Approaches to Predictive Maintenance: A Systematic Literature Survey on Diagnostics and Prognostics. *Journal of Manufacturing Systems*. 2020. vol. 56. pp. 539–557.
 18. Zhang P., Li T., Wang G., Luo C. Multi-source Information Fusion Based on Rough Set Theory: A Review. *Information Fusion*. 2021. vol. 68. pp. 85–117.
 19. Li Z., Tian L., Jiang Q., Yan X. Fault Diagnostic Method Based on Deep Learning and Multimodel Feature Fusion for Complex Industrial Processes. *Industrial & Engineering Chemistry Research*. 2020. vol. 59(40). pp. 18061–18069.
 20. Shahraki A.F., Yadav O.P., Liao H. A Review on Degradation Modelling and Its Engineering Applications. *International Journal of Performability Engineering*. 2017. vol. 13. pp. 299–314.
 21. Yakimov V.L. [The Forecasting of Launch Complex Technical State Parameters Using Neural Networks]. *Izvestiia vysshikh uchebnykh zavedenii. Priborostroenie – Proceedings of higher educational institutions. Instrumentation*. 2006. vol. 7. pp. 7–10. (In Russ.).
 22. Marushko E.E., Doudkin A.A. Ensembles of Neural Networks for Forecasting of Time Series of Spacecraft Telemetry. *Optical Memory and Neural Networks*. 2017. vol. 26. no. 1. pp. 47–54.
 23. Emel'ianova Iu.G., Fralenko V.P. [Methods of Information Cognitive-Graphic Presentation for Complex Technical Systems Effective Monitoring]. *Programmnye sistemy: teoriia i prilozheniia – Software Systems: Theory and Applications*. 2018. vol. 9. no. 4(39). pp. 117–158. (In Russ.).
 24. Mal'tsev G.N., Nazarov A.V., Yakimov V.L. [Training Sample Entropy Maximizing Algorithm and Its Use in Synthesis of Predictive Model Nonlinear Dynamic System Discrete States]. *Informatsionno-upravliaiushchie sistemy – Information management systems*. 2015. no. 2. pp. 57–66. (In Russ.).
 25. Preece A. Asking «Why» in AI: Explainability of Intelligent Systems – Perspectives and Challenges. *Intelligent Systems in Accounting, Finance and Management*. 2018. vol. 25. pp. 63–72.
 26. Marcot B.G., Penman T.D. Advances in Bayesian Network Modelling: Integration of Modelling Technologies. *Environmental Modelling and Software*. 2019. vol. 111. pp. 386–393.
 27. Chen J., Pi D., Wu Z., Zhao X. Imbalanced Satellite Telemetry Data Anomaly Detection Model Based on Bayesian LSTM. *Acta Astronautica*. 2021. vol. 180. pp. 232–242.
 28. Maslennikov E.D., Sulimov V.B. [Bayesian Trust Network Predictions: Algorithm and Software Implementation]. *Vychislitel'nye metody i programirovanie – Computational Methods and Programming*. 2010. vol. 11. no. 4. pp. 94–107. (In Russ.).
 29. Musina V.F. [Bayesian Networks of Trust as Probabilistic Graphical Model for Assessing Medical Risks]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. Issue 1(24). pp. 135–151. (In Russ.).
 30. Gao X., Guo Z., Ren H., Yang Y. Learning Bayesian network parameters via minimax algorithm. *International Journal of Approximate Reasoning*. 2019. vol. 108. pp. 62–75.

31. Toropova A.V. [Approaches to Diagnosing Data Consistency in Bayesian Trust Networks]. Trudy SPIIRAN – SPIIRAS Proceedings. 2015. Issue 6(43). pp. 156–178. (In Russ.).
32. Smirnov V.A. [Troubleshooting On-Board Control Systems During Acceptance Control]. Informatsionno-upravliaiushchie sistemy – Information management systems. 2013. no. 2. pp. 24–28. (In Russ.).
33. Larin V.P., Shelest D.K. [Formation of Information Support for Reliability of Onboard Equipment at Design Stage]. Informatsionno-upravliaiushchie sistemy – Information management systems. 2012. no. 4. pp. 93–97. (In Russ.).
34. Akhmetov R.N. [Methods and Models of Autonomous Control Over Earth Remote Sensing Spacecraft Survivability]. Vestnik Samarskogo gosudarstvennogo aerokosmicheskogo universiteta – Bulletin of the Samara State Aerospace University. 2008. no. 3. pp. 194–210. (In Russ.).
35. Kresoja M., Luzanin Z., Stojkovska I. Adaptive Stochastic Approximation Algorithm. Numerical Algorithms. 2017. vol. 76. pp. 917–937.
36. Granichin O.N., Izmakova O.A. [A Randomized Stochastic Approximation Algorithm in Self-Learning Problem]. Avtomatika i telemekhanika – Automation and telemechanics. 2005. no. 8. pp. 52–63. (In Russ.).
37. Wangab D., Tiana L. Parametric Methods for Confidence Interval Estimation of Overlap Coefficients. Computational Statistics & Data Analysis. 2017. vol. 106. pp. 12–26.
38. Mal'tsev G.N., Yakimov V.L., Solov'ev S.V., Lebedeva N.V. [Primary Processing of Telemetry Information Using Dynamic Models of Parameter Changes and Partial Nonlinear Filtering]. Informatsionno-upravliaiushchie sistemy – Information management systems. 2018. no. 5. pp. 22–34. (In Russ.).
39. Andoga R., Fozo L., Judicak J., Breda R. Intelligent Situational Control of Small Turbojet Engines. International Journal of Aerospace Engineering. 2018. pp. 1–16.
40. Gorod A., Ireland V., Gunawan I., Hallo L. Evolving Toolbox for Complex Project Management. Boca Raton: Auerbach Publications, 2019. 570 p.

Г.Б. МАРШАЛКО, Ю.А. ТРУФАНОВА
**ПОЛИНОМИАЛЬНЫЕ АППРОКСИМАЦИИ НЕКОТОРЫХ
ФУНКЦИЙ АКТИВАЦИИ НЕЙРОННЫХ СЕТЕЙ**

Маршалко Г.Б., Труфанова Ю.А. Полиномиальные аппроксимации некоторых функций активации нейронных сетей.

Аннотация. Активное внедрение систем машинного обучения ставит актуальную задачу обеспечения их защиты от различных типов атак, направленных на нарушение свойств конфиденциальности, целостности и доступности как обрабатываемых данных, так и обучаемых моделей. Одним из перспективных направлений защиты является разработка конфиденциальных систем машинного обучения, использующих гомоморфные схемы шифрования для защиты моделей и данных. Однако такие схемы могут обрабатывать только полиномиальные функции, что в свою очередь ставит задачу построения полиномиальных аппроксимаций используемых в нейросетевых моделях нелинейных функций. Целью настоящей работы является построение наиболее точных аппроксимаций некоторых широко используемых функций активаций нейронных сетей, а именно ReLU, логистического сигмоида и гиперболического тангенса, при ограничениях на степень аппроксимирующего полинома, а также оценка влияния точности такой аппроксимации на результат работы нейронной сети в целом. В отличие от опубликованных ранее работ рассматриваются и сравниваются различные способы построения аппроксимирующих полиномов, вводятся метрики точности приближения, приводится конкретный вид аппроксимирующих полиномов, а также соответствующие значения точности приближения. Проводится сравнение с аппроксимациями, приведенными в опубликованных ранее работах. В заключение для простейшей нейронной сети экспериментально оценено влияние точности приближения аппроксимирующего полинома на величину отклонения значений выходных нейронов такой сети от соответствующих значений выходных нейронов исходной сети. Результаты показывают, что для функции ReLU наилучшее приближение может быть получено с помощью численного метода, а для логистического сигмоида и гиперболического тангенса – с помощью полиномов Чебышева. При этом наилучшее приближение из трех рассмотренных функций получено для функции ReLU. Полученные результаты в дальнейшем могут быть использованы при построении аппроксимаций функций активации в конфиденциальных системах машинного обучения.

Ключевые слова: функция активации, ReLU, гиперболический тангенс, логистический сигмоид, гомоморфное шифрование, BGV, CKKS, нейронная сеть, полиномиальная аппроксимация, конфиденциальное машинное обучение.

1. Введение. Одним из способов защиты систем машинного обучения от широкого спектра существующих атак [1] является построение систем конфиденциального машинного обучения с использованием схем гомоморфного шифрования [2–4]. Примерами используемых схем шифрования являются BGV [5], CKKS [6] и др.

Необходимым условием реализации вычислений при таком методе защиты является замена нелинейных функций активации нейронов их полиномиальными представлениями. Если степень нелинейности таких представлений будет относительно мала, то применение гомоморфных

схем, таких, например, как *BGV* позволит обеспечить гарантированную защиту данных и/или моделей при сравнительно небольшом снижении скорости работы. Это связано с тем, что современные схемы гомоморфного шифрования основаны на сложности решения задачи обучения с ошибками (LWE) или обучения с ошибками в кольце (RLWE) и используют при формировании шифртекста шумовые последовательности. При выполнении арифметических операций «амплитуда» шума увеличивается и, начиная с некоторого момента, корректное расшифрование становится невозможным. Особенно быстро шум накапливается при умножении. Снизить уровень шума можно с помощью т.н. операции перезагрузки, однако она выполняется достаточно медленно, что существенно снижает скорость шифрования. В этой связи важно ограничивать степень гомоморфно вычисляемого полинома. С другой стороны, для сохранения свойств обученных моделей машинного обучения нам необходимо использовать достаточно точные аппроксимации исходных функций, что, очевидно, накладывает ограничения на минимально допустимое значение степени полинома.

Исследователи в области применения гомоморфного шифрования к нейронным сетям рассматривали и другой способ обработки нелинейных функций. Он заключается в предварительном вычислении значений функции активации для набора дискретных значений и сохранении этих результатов в поисковую таблицу [7]. Однако, в этом случае качество работы нейронной сети напрямую зависит от числа записей в поисковой таблице, то есть данный подход требует больших затрат по памяти. Поэтому ограничимся рассмотрением полиномиальных аппроксимаций для функций активации.

В доступной литературе по данной тематике практически отсутствуют подробные описания методов построения полиномиальных аппроксимаций функций активации нейронных сетей, конкретного вида получающихся полиномов и обоснования точности аппроксимации.

Одной из первых работ, в которой были описаны принципы применения гомоморфных схем шифрования в задачах машинного обучения, была работа исследователей Microsoft [2].

В некоторых работах, в частности [2, 8], было предложено заменить функцию активации *ReLU* полиномом x^2 . Однако, в работе [9] было показано, что такая аппроксимация хорошо работает только на нейронных сетях с небольшим числом слоев. Для глубоких нейронных сетей качество работы при замене функций активации *ReLU* на $x^2 + x$ выше, чем при замене на x^2 , хотя и проигрывает оригинальной сети.

В работе [3] авторы кратко описали подходы, которые могут быть применены для аппроксимации таких широко используемых функций активации как *ReLU*, *логистический сигмоид* и *гиперболический тангенс*. Однако конкретный вид аппроксимирующих полиномов приведен не был.

В репозитории GitHub представлен проект [10], реализующий описанный в работе [3] подход. Авторы проекта представили полиномы, полученные при помощи четырех методов аппроксимации для функций активации *ReLU* и *гиперболический тангенс* степени, не превышающей 4, причем ни один из графиков полученных функций не соответствует результатам, представленным в работе [3].

В работе [11] представлены результаты исследований по оценке близости полиномиальных приближений нейронных сетей к исходным в зависимости от параметра точности α . Авторы рассматривают функцию ошибки аппроксимированной модели по отношению к исходной, показывая, что она может быть сделана сколь угодно малой: $|F(x) - F_\alpha(x)| \leq Const \cdot 2^{-\alpha}$. Здесь $F(x)$ – это функция, описывающая исходную нейронную сеть, $F_\alpha(x)$ – приближение с точностью α , в котором степень аппроксимирующих многочленов определенным образом зависит от данного параметра, $Const$ – некоторая константа. В этой работе авторы в явном виде приводят вид аппроксимирующих полиномов, позволяющих получить существенно более точные нейросетевые модели по сравнению с работой [3]. Вместе с тем полученные ими полиномы имеют степень не менее 7, что, как было указано выше, существенно замедляет скорость работы защищенной модели.

Указанные выше работы не раскрывают деталей процесса выработки приведенных в них аппроксимирующих полиномов. В связи с этим вопрос о выборе наилучшего метода приближения, равно как и точности получаемых приближений, остается открытым. Настоящая работа посвящена исследованию различных способов приближения нелинейных функций активации, оценке взаимосвязи точности и степени нелинейности таких приближений и получению явных представлений таких приближений.

2. Основные сведения и обозначения.

2.1. Функции активации. Рассматриваются следующие наиболее часто используемые при построении нейронных сетей функции активации (Рис. 1):

$$\begin{aligned} - \text{Rectified linear units (ReLU)}: & \quad ReLU(x) = \max(0, x); \\ - \text{Логистический сигмоид (Sigmoid)}: & \quad \sigma(x) = \frac{1}{1+e^{-x}}; \\ - \text{Гиперболический тангенс}: & \quad \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \end{aligned}$$

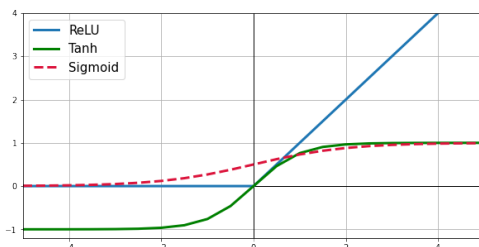


Рис. 1. Исследуемые функции активации

2.2. Методы аппроксимации. Пусть $f(x)$ – функция, которую мы хотим аппроксимировать полиномом $p(x)$ степени, не выше N . Функция $f(x)$ определена на отрезке $[a, b]$. Точки $x_i \in [a, b]$ – узлы интерполяции ($i = 1, \dots, m; m \geq N$).

Будем исследовать следующие методы аппроксимации:

1. *Ряд Тейлора в окрестностях точки $z \in [a, b]$*

Разложение строится численным методом. Основная идея заключается в построении полинома, приближающего исходную функцию на заданном отрезке, методом Круга [12], и разложении этой функции в ряд Тейлора в окрестностях заданной точки.

2. *Численный метод*

Коэффициенты полинома $p(x)$ степени N подбираются при помощи логистической регрессии таким образом, чтобы минимизировать среднюю квадратичную ошибку (mean squared error - MSE) в узлах интерполяции $\{x_i, f(x_i)\}_{i=1}^m$.

3. *Аппроксимация полиномами Чебышёва*

Аппроксимация функции $f(x)$ на отрезке $[a, b]$ строится в форме полинома следующего вида [13]:

$$f(x) \simeq \sum_{i=0}^N \alpha_i T_i \left(2 \frac{x-a}{b-a} - 1 \right),$$

где $T_i(x) = \cos(i \arccos(x))$, $x \in [-1, 1]$ – полиномы Чебышёва, α_i – соответствующие коэффициенты. Для T_i справедливы следующие равенства:

$$T_0(x) = 1, T_1(x) = x,$$

$$T_{i+1}(x) = 2xT_i(x) - T_{i-1}(x), i = 2, 3, \dots$$

4. Подход, основанный на аппроксимации производной

Пусть $p_1(x)$ – полиномиальное приближение для производной $f'(x)$ на отрезке $[a, b]$, полученное одним из методов, описанных ранее (п. 1 – 3), $\deg p_1(x) = (N - 1)$. Тогда аппроксимацией функции $f(x)$ на отрезке $[a, b]$ является полином $p(x)$ степени N [3]:

$$p = \int p_1(x) dx.$$

Будем рассматривать следующие способы выбора узлов интерполяции $\{x_i\}_{i=1}^m$:

- равномерное распределение на отрезке $x_i \sim U(a, b)$ (Рис. 2);

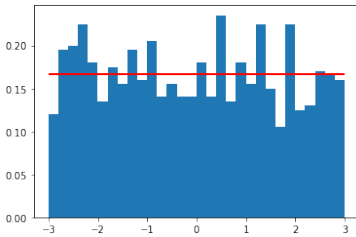


Рис. 2. $x_i \sim U(-3, 3)$

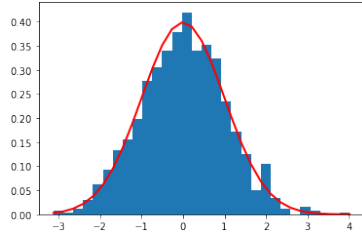


Рис. 3. $x_i \sim N(0, 1)P(x_i \in [-3, 3]) \approx 0,99$

- нормальное распределение $x_i \sim N(\mu, \sigma^2)$ (Рис. 3), где

$$\mu = [(a + b)/2], \sigma = [(b - a)/6] \Rightarrow P(x_i \in [a, b]) \approx 0,9973;$$

$$- x_i = (r_i + 1) \frac{b-a}{2} + a, \quad k = 1, \dots, m,$$

где $m > N$ – число узлов интерполяции (Рис. 4),

r_i – корни многочленов Чебышёва степени не выше N :

$$r_i = -\cos \left(\frac{2i - 1}{2m} \pi \right).$$

По графикам видно, что концентрация узлов интерполяции:

- примерно одинакова на всем отрезке при равномерном распределении,
- максимальна в середине отрезка и уменьшается к его концам при нормальном распределении,

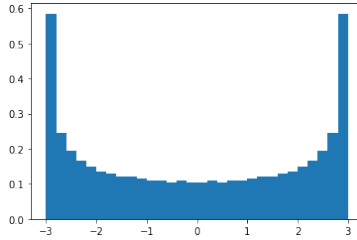


Рис. 4. x_i – корни многочленов Чебышёва, отображённые на отрезок $[-3, 3]$

– максимальна на краях отрезка и уменьшается к середине для корней многочленов Чебышёва.

Следовательно, можно ожидать, что при равномерном распределении точек полином будет стремиться приблизить исходную функцию одинаково хорошо на всем отрезке, при нормальном распределении – в середине отрезка, для корней многочленов Чебышёва – на концах отрезка.

Отметим, что для аппроксимации функции через разложение в ряд Тейлора узлы интерполяции не используются. Нужно знать значение функции и ее производных лишь в одной точке, в окрестностях которой выполняется аппроксимация.

3. Вводимые ограничения.

3.1. Способы оценки качества аппроксимации. Существуют различные способы оценки качества аппроксимации, такие как:

– средняя абсолютная ошибка

$$MAE(f, p) = \frac{1}{m} \sum_{i=1}^m |f(x_i) - p(x_i)|,$$

– средняя квадратичная ошибка

$$MSE(f, p) = \frac{1}{m} \sum_{i=1}^m (f(x_i) - p(x_i))^2,$$

– максимальное значение разности

$$\max_{x_i} |f(x_i) - p(x_i)|$$

$i=1, \dots, m$

и др.

Оценки MAE и MSE имеют схожий смысл, поэтому в данной работе будут рассматриваться MSE и max_loss , которые выявляют различные типы отклонений значений исследуемых функций.

В качестве единой метрики для сравнения точности аппроксимации будем использовать

$$\rho(f, p) = \sqrt{MSE(f, p)^2 + max_loss(f, p)^2}.$$

Данная метрика может рассматриваться как аналог естественной евклидовой метрики в двумерном пространстве, где метрики MSE и max_loss определяют расстояние по соответствующим координатам.

3.2. Выбор отрезка для аппроксимации. Рассмотрим поведение функций аппроксимации (Рис. 1):

– *ReLU* – кусочно-полиномиальная на отрезках $(-\infty, 0]$, $[0, \infty)$, принимает значения из отрезка $[0, \infty)$;

– *Sigmoid* – монотонно возрастает, точка перегиба в 0, $\sigma(x) \in (0, 1)$, $\sigma(x) \xrightarrow{x \rightarrow -\infty} 0$, $\sigma(x) \xrightarrow{x \rightarrow +\infty} 1$.

– *Tanh* – монотонно возрастает, точка перегиба в 0, $tanh(x) \in (-1, 1)$, $tanh(x) \xrightarrow{x \rightarrow -\infty} -1$, $tanh(x) \xrightarrow{x \rightarrow +\infty} 1$.

Исходя из вышесказанного, очевидно, что необходимо выбирать отрезок аппроксимации, симметричный относительно нуля, вида $[-l, l]$, $l \neq 0$. В работе [3, 10] использовалось значение $l = 200$. Мы также используем отрезок аппроксимации $[-200, 200]$ в целях сравнения результатов.

Поскольку множества значений рассматриваемых функций активации различны, для корректного сравнения качества приближения, которое показывают рассматриваемые методы аппроксимации, нормализуем значения функций и полиномов с помощью проекции на отрезок $[0, 1]$:

$$y_{normalized} = \frac{y - y_{min}}{y_{max} - y_{min}} :$$

– *ReLU*(x) : $y_{min} = 0, y_{max} = 200$;

– $\sigma(x)$: $y_{min} = 0, y_{max} = 1$;

– *tanh*(x) : $y_{min} = -1, y_{max} = 1$.

После такой замены значения MSE и max_loss изменятся следующим образом:

$$MSE_{normalized} = \frac{MSE}{(y_{max} - y_{min})^2},$$

$$\max_loss_{normalized} = \frac{\max_loss}{(y_{max} - y_{min})}.$$

Далее в таблицах 1, 2, 3, 4, 5 нижние индексы $MSE_{normalized}$ и $\max_loss_{normalized}$ опущены для краткости.

3.3. Выбор степени полинома. Поскольку, как было отмечено во введении, существующие схемы гомоморфного шифрования ограничивают допустимое количество операций умножения, разработчикам необходимо по возможности ограничивать степень аппроксимирующего полинома. Вместе с тем, конкретные значения степеней будут зависеть от вида нейронной сети и параметров используемой схемы гомоморфного шифрования. В связи с этим далее будут рассматриваться полиномы степеней 2, 3, 4 аналогично [3, 10] с тем, чтобы иметь возможность сравнения полученных результатов.

4. Полученные результаты. Далее будут использованы следующие обозначения:

- MSE – нормализованная среднеквадратичная ошибка приближения исходной функции полиномом,
- \max_loss – нормализованное максимальное значение разности между полученным полиномом и исходной функцией.

Необходимо сделать некоторые замечания относительно построения приближений для функции $ReLU$.

Функция активации $ReLU$ недифференцируема в нуле, поэтому не имеет разложения в ряд Тейлора в окрестностях этой точки. Однако, поскольку для построения разложения в ряд Тейлора используется численный метод, это ограничение можно обойти.

Производной функции $ReLU$ является функция Хевисайда:

$$ReLU'(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}.$$

Эта функция разрывна и недифференцируема в нуле. Поэтому авторами работы [3] было предложено вместо нее аппроксимировать функцию логистического сигмоида, которая имеет асимптоты $y = 0$ при $x \rightarrow -\infty$ и $y = 1$ при $x \rightarrow +\infty$, то есть приближает функцию Хевисайда.

4.1. Полученные полиномы. В таблицах 1, 2, 3 далее приведены полиномиальные аппроксимации функций $ReLU$, логистического сигмоида и гиперболического тангенса, соответственно. Для аппроксимации производной использовался метод аппроксимации многочленами Чебышёва, так как он проявил себя наилучшим образом среди рассмотренных методов.

Таблица 1. Полиномиальные аппроксимации для $ReLU$

Метод	Полиномы	MSE	max_loss	ρ
Численный метод. $x_i \sim U(-200, 200)$	$0.00236x^2 + 0.5006x + 18.5 - 2.834 \cdot 10^{-7}x^3 + 0.002358x^2 + 0.5074x + 18.47$	0.0013 0.0013	0.0925 0.0924	0.0925 0.0924
	$-5.178 \cdot 10^{-8}x^4 - 2.154 \cdot 10^{-8}x^3 + 0.004124x^2 + 0.5021x + 11.49$	0.0003	0.0575	0.0575
Численный метод. $x_i \sim N(0, 67^2)$	$0.0030x^2 + 0.4915x + 13.25 - 2.011 \cdot 10^{-6}x^3 + 0.003051x^2 + 0.5175x + 13.07$	0.0032 0.004	0.1734 0.2385	0.1735 0.2385
	$-5.114 \cdot 10^{-8}x^4 + 1.103 \cdot 10^{-6}x^3 + 0.004391x^2 + 0.4841x + 10.1$	0.0008	0.0614	0.0614
Численный метод. x_i – корни многочленов Чебышёва.	$0.002122x^2 + 0.5x + 21.22 - 7.846 \cdot 10^{-21}x^3 + 0.002122x^2 + 0.5x + 21.22$	0.0015 0.0015	0.1061 0.1061	0.1061 0.1061
	$-4.244 \cdot 10^{-8}x^4 + 6.431 \cdot 10^{-21}x^3 + 0.00382x^2 + 0.5x + 12.73$	0.0004	0.0637	0.0637
Ряд Тейлора	$0.001472x^2 + 0.5x + 40 - 4.466 \cdot 10^{-6}x^3 + 0.003028x^2 + 0.634x + 8.965 \cdot 10^{-15}$	0.0063 0.0063	0.2 0.1503	0.2 0.1504
	$-1.03 \cdot 10^{-8}x^4 + 3.6 \cdot 10^{-21}x^3 + 0.002091x^2 + 0.5x + 28.57$	0.0025	0.1429	0.1429
Полиномы Чебышёва	$0.002122x^2 + 0.5x + 21.22 - 3.217 \cdot 10^{-21}x^3 + 0.002122x^2 + 0.5x + 21.22$	0.0015 0.0015	0.1061 0.1061	0.1061 0.1061
	$-4.244 \cdot 10^{-8}x^4 - 9.181e \cdot 10^{-21}x^3 + 0.00382x^2 + 0.5x + 12.73$	0.0004	0.0637	0.0637
Аппроксимация производной (функция Хевисайда)	$0.001592x^2 + 0.5x + 36.34 - 1.017 \cdot 10^{-21}x^3 + 0.001592x^2 + 0.5x + 36.34$	0.0047 0.0047	0.1817 0.1817	0.1818 0.1818
	$-2.653 \cdot 10^{-8}x^4 - 7.138 \cdot 10^{-21}x^3 + 0.003183x^2 + 0.5x + 15.12$	0.0006	0.0756	0.0756
Аппроксимация производной (логистический сигмоид)	$0.001591x^2 + 0.5x + 36.34 - 3.311 \cdot 10^{-21}x^3 + 0.001591x^2 + 0.5x + 36.34$	0.0047 0.0047	0.1817 0.1817	0.1818 0.1818
	$-2.652 \cdot 10^{-8}x^4 - 7.138 \cdot 10^{-21}x^3 + 0.003182x^2 + 0.5x + 15.13$	0.0006	0.0756	0.0756

Таблица 2. Полиномиальные аппроксимации для логистического сигмоида

Метод	Полиномы	MSE	max_loss	ρ
Численный метод. $x_i \sim U(-200, 200)$	$-5.352 \cdot 10^{-7}x^2 + 0.003761x + 0.5167$	0.0602	0.4917	0.4953
	$-1.378 \cdot 10^{-7}x^3 - 3.097 \cdot 10^{-7}x^2 + 0.007042x + 0.5127$	0.0328	0.4679	0.4691
	$5.441 \cdot 10^{-11}x^4 - 1.378 \cdot 10^{-7}x^3 - 2.166 \cdot 10^{-6}x^2 + 0.007041x + 0.5201$	0.0329	0.4754	0.4765
Численный метод. $x_i \sim N(0, 67^2)$	$-7.977 \cdot 10^{-7}x^2 + 0.006058x + 0.4916$	0.1333	0.7518	0.7636
	$-2.395 \cdot 10^{-7}x^3 + 1.031 \cdot 10^{-6}x^2 + 0.00917x + 0.487$	0.0512	0.6101	0.6122
	$2.507 \cdot 10^{-10}x^4 - 2.456 \cdot 10^{-7}x^3 - 5.33 \cdot 10^{-6}x^2 + 0.0093x + 0.5$	0.0565	0.8034	0.8053
Численный метод. x_i – корни многочленов Чебышёва.	$2.292 \cdot 10^{-20}x^2 + 0.003183x + 0.5$	0.0641	0.4784	0.4827
	$-1.061 \cdot 10^{-7}x^3 - 9.173 \cdot 10^{-21}x^2 + 0.006365x + 0.5$	0.0340	0.4594	0.4606
	$9.441 \cdot 10^{-25}x^4 - 1.061 \cdot 10^{-7}x^3 - 2.817 \cdot 10^{-20}x^2 + 0.006365x + 0.5$	0.0340	0.4594	0.4606
Ряд Тейлора	$-7.5 \cdot 10^{-6}x^2 + 0.004472x + 0.6$	0.0756	0.5944	0.5992
	$-2.437 \cdot 10^{-8}x^3 - 2.054 \cdot 10^{-19}x^2 + 0.003363x + 0.5$	0.0634	0.4774	0.4815
	$2.232 \cdot 10^{-10}x^4 - 9.958 \cdot 10^{-8}x^3 - 1.071 \cdot 10^{-5}x^2 + 0.006377x + 0.5714$	0.0374	0.5303	0.5316
Полиномы Чебышёва	$-9.932 \cdot 10^{-21}x^2 + 0.003183x + 0.5$	0.0641	0.4784	0.4827
	$-1.061 \cdot 10^{-7}x^3 - 2.141 \cdot 10^{-20}x^2 + 0.006365x + 0.5$	0.0340	0.4594	0.4606
	$-1.248 \cdot 10^{-25}x^4 - 1.061 \cdot 10^{-7}x^3 - 1.901 \cdot 10^{-20}x^2 + 0.006365x + 0.5$	0.0340	0.4594	0.4606
Аппроксимация производной	$-5.818 \cdot 10^{22}x^2 + 0.001592x + 0.3183$	0.1547	0.6697	0.6872
	$-5.305 \cdot 10^{-8}x^3 + 0.004774x + 0.5305$	0.0445	0.4994	0.5014
	$2.182 \cdot 10^{-26}x^4 - 5.305 \cdot 10^{-8}x^3 - 3.103 \cdot 10^{-21}x^2 + 0.004774x + 0.5305$	0.0445	0.4994	0.5014

В таблицах 4 и 5 мы приведем явный вид и оценки точности аппроксимации полиномов из проекта CryptoDL [3, 10].

Таблица 3. Полиномиальные аппроксимации для гиперболического тангенса

Метод	Полиномы	MSE	max_loss	ρ
Численный метод. $x_i \sim U(-200, 200)$	$7.093 \cdot 10^{-8}x^2 +$ $0.007516x - 0.01842$	0.0612	0.4938	0.4976
	$-2.775 \cdot 10^{-7}x^3 - 1.734 \cdot$ $10^{-7}x^2 + 0.01418x +$ 0.002386	0.0338	0.4725	0.4737
	$7.749 \cdot 10^{-11}x^4 - 2.776 \cdot$ $10^{-7}x^3 - 2.809 \cdot 10^{-6}x^2 +$ $0.01418x + 0.01252$	0.0338	0.4776	0.4788
Численный метод. $x_i \sim N(0, 67^2)$	$-7.288 \cdot 10^{-7}x^2 +$ $0.01221x - 0.004985$	0.1369	0.7378	0.7504
	$-4.893 \cdot 10^{-7}x^3 + 5.002 \cdot$ $10^{-7}x^2 + 0.01845x -$ 0.01014	0.0545	0.617	0.6194
	$5.371 \cdot 10^{-10}x^4 - 5.007 \cdot$ $10^{-7}x^3 - 1.295 \cdot 10^{-5}x^2 +$ $0.01861x + 0.01763$	0.0596	0.8218	0.824
Численный метод. x_i – корни многочленов Чебышёва.	$-5.393 \cdot 10^{-21}x^2 +$ $0.006366x$	0.0651	0.4869	0.4913
	$-2.122 \cdot 10^{-7}x^3 + 2.024 \cdot$ $10^{-21}x^2 + 0.01273x + 6.64 \cdot$ 10^{-17}	0.0350	0.4742	0.4755
	$3.092 \cdot 10^{-24}x^4 - 2.122 \cdot$ $10^{-7}x^3 - 1.32 \cdot 10^{-19}x^2 +$ $0.01273x + 9.531 \cdot 10^{-16}$	0.0350	0.4742	0.4755
Ряд Тейлора	$-1.5 \cdot 10^{-5}x^2 + 0.008944x +$ 0.2	0.0766	0.5944	0.5993
	$-4.874 \cdot 10^{-8}x^3 - 8.165 \cdot$ $10^{-19}x^2 + 0.006726x +$ $2.429 \cdot 10^{-14}$	0.0645	0.4862	0.4905
	$4.464 \cdot 10^{-10}x^4 - 1.992 \cdot$ $10^{-7}x^3 - 2.143 \cdot 10^{-5}x^2 +$ $0.01275x + 0.1429$	0.0384	0.5454	0.5467
Полиномы Чебышёва	$-1.835 \cdot 10^{-20}x^2 +$ $0.006366x + 3.309 \cdot 10^{-16}$	0.0651	0.4869	0.4912
	$-2.122 \cdot 10^{-7}x^3 - 1.835 \cdot$ $10^{-20}x^2 + 0.01273x +$ $3.309 \cdot 10^{-16}$	0.0350	0.4742	0.4755
	$1.048 \cdot 10^{-24}x^4 - 2.122 \cdot$ $10^{-7}x^3 - 4.442 \cdot 10^{-20}x^2 +$ $0.01273x + 8.387 \cdot 10^{-17}$	0.0350	0.4742	0.4755
Аппроксимация производной	$1.115 \cdot 10^{-21}x^2 + 0.003183x$	0.1227	0.4933	0.5083
	$-1.061 \cdot 10^{-7}x^3 +$ $0.009549x$	0.0446	0.4806	0.4826
	$5.091 \cdot 10^{-26}x^4 - 1.061 \cdot$ $10^{-7}x^3 - 6.206 \cdot 10^{-21}x^2 +$ $0.009549x$	0.0446	0.4806	0.4826

Таблица 4. Полиномиальные аппроксимации для ReLU из проекта [3, 10]

Метод	Полиномы	MSE	max_loss	ρ
Полиномы Лежандра	$0.0021x^2 + 0.5x + 21.22$	0.0015	0.1061	0.1061
	$-3.517 \cdot 10^{-16}x^3 + 0.0021x^2 + 0.5x + 21.22$	0.0015	0.1061	0.1061
	$-4.244 \cdot 10^{-8}x^4 + -3.517 \cdot 10^{-16}x^3 + 0.0038x^2 + 0.5x + 12.732$	0.0004	0.0637	0.0637
Численный метод	$0.0016x^2 + 0.5x + 38.557$	0.0057	0.1928	0.1929
	$-1.6093 \cdot 10^{-14}x^3 + 0.0016x^2 + 0.5x + 38.557$	0.0057	0.1928	0.1929
	$-1.73 \cdot 10^{-8}x^4 - 3.61 \cdot 10^{-14}x^3 + 0.0025x^2 + 0.5x + 27.453$	0.0021	0.1373	0.1373
Полиномы Чебышёва	$0.00234x^2 + 0.5x + 18.75$	0.0013	0.0938	0.0938
	$9.363 \cdot 10^{-20}x^3 + 0.00234x^2 + 0.5x + 18.75$	0.0013	0.0938	0.0938
	$-5.127 \cdot 10^{-8}x^4 + 9.363 \cdot 10^{-20}x^3 + 0.0041x^2 + 0.5x + 11.719$	0.0003	0.0586	0.0586
Аппроксимация производной (сигмоид)	$0.0013x^2 + 0.5x$	0.0106	0.2344	0.2346
	$-9.677 \cdot 10^{-21}x^3 + 0.002x^2 + 0.5x$	0.0043	0.2344	0.2346

В сводных таблицах 6 и 7 приведены наилучший метод аппроксимации и соответствующее значение метрики ρ для каждой рассмотренной степени аппроксимирующего полинома и каждой рассмотренной функции активации, полученные в данном исследовании и использованные в проекте [3, 10], соответственно. В целом, полученные результаты согласуются в смысле порядка ошибки аппроксимации. Вместе с тем, в настоящем исследовании удалось получить несколько более точную аппроксимацию для функции $ReLU(x)$, чем в проекте [3, 10].

Полученные аппроксимации для логистического сигмоида существенно проигрывают в точности функции $ReLU$, однако оказались несколько лучше, чем аппроксимации для гиперболического тангенса.

Как было отмечено во введении, задача выбора аппроксимирующего полинома является по меньшей мере двухпараметрической (минимизация и ошибки аппроксимации, и степени полинома). Поэтому представляет интерес сравнение ошибок при разных степенях. Для каждой функции активации найдем полиномы степеней 2,3,4, дающие наилучшее приближе-

Таблица 5. Полиномиальные аппроксимации для гиперболического тангенса из проекта [3, 10]

Метод	Полиномы	MSE	max_loss	ρ
Полиномы Лежандра	$0.00637x$	0.0651	0.4869	0.4912
	$-2.1219 \cdot 10^{-7}x^3 + 0.0127x$	0.035	0.4742	0.4755
	$-2.1219 \cdot 10^{-7}x^3 + 0.0127x$	0.035	0.4742	0.4755
Численный метод	$0.00587x$	0,0696	0.4879	0.4928
	$-1.2015 \cdot 10^{-7}x^3 + 0.0095x$	0.0445	0.4806	0.4827
	$-1.2015 \cdot 10^{-7}x^3 + 0.0095x$	0.0445	0.4806	0.4827
Полиномы Чебышёва	$0.00745x$	0,0611	0.4847	0.4885
	$-2.734 \cdot 10^{-7}x^3 + 0.0141x$	0.0338	0.4716	0.4728
	$-2.734 \cdot 10^{-7}x^3 + 0.0141x$	0.0338	0.4716	0.4728

Таблица 6. Сводная таблица аппроксимаций функций активации

Функция	Степень полинома			Лучший результат
	2	3	4	
<i>ReLU</i>	Численный метод. $x_i \sim U(-200, 200)$, $\rho = 0.0925$	Численный метод. $x_i \sim U(-200, 200)$, $\rho = 0.0924$	Численный метод. $x_i \sim U(-200, 200)$, $\rho = 0.0575$	Численный метод. $x_i \sim U(-200, 200)$, $deg = 4$, $\rho = 0.0575$
<i>Sigmoid</i>	Полиномы Чебышёва, $\rho = 0.4827$	Полиномы Чебышёва, $\rho = 0.4606$	Полиномы Чебышёва, $\rho = 0.4606$	Полиномы Чебышёва, $deg = 3$, $\rho = 0.4606$
<i>Tanh</i>	Полиномы Чебышёва, $\rho = 0.4912$	Численный метод. $x_i \sim U(-200, 200)$ $\rho = 0.4737$	Полиномы Чебышёва, $\rho = 0.4755$	Численный метод. $x_i \sim U(-200, 200)$, $deg = 3$, $\rho = 0.4737$

ние. Для каждого из них рассчитаем ошибку аппроксимации и вычислим модули их попарных разностей. Полученные результаты представлены в таблице 8.

Можно видеть, что для:

– функции *ReLU* использование полинома третьей степени практически не дает выигрыша по сравнению с использованием полинома

Таблица 7. Сводная таблица аппроксимаций функций активации из проекта [3, 10]

Функция	Степень полинома			Лучший результат
	2	3	4	
<i>ReLU</i>	Полиномы Чебышёва, $\rho = 0.0938$	Полиномы Чебышёва, $\rho = 0.0938$	Полиномы Чебышёва, $\rho = 0.0586$	Полиномы Чебышёва, $deg = 4,$ $\rho = 0.0586$
<i>Tanh</i>	Полиномы Чебышёва, $\rho = 0.4885$	Полиномы Чебышёва, $\rho = 0.4728$	Полиномы Чебышёва, $\rho = 0.4728$	Полиномы Чебышёва, $deg = 3,$ $\rho = 0,4728$

Таблица 8. Таблица модуля разности ошибки аппроксимации для наилучших аппроксимирующих многочленов различных степеней

Функция	Степени полиномов		
	2 – 3	2 – 4	3 – 4
<i>ReLU</i>	0.001	0.035	0.034
<i>Sigmoid</i>	0.0221	0.0221	0
<i>Tanh</i>	0.0239	0.0188	0.0051

второй степени, хотя они уступают по точности полиному четвертой степени;

- логистического сигмоида, отсутствует разница между ошибками аппроксимаций 3 и 4 степеней;
- гиперболического тангенса наилучшим выбором является использование полинома третьей степени.

Эти результаты целесообразно использовать для оптимизации реализаций гомоморфных вычислений.

5. Оценка ошибки нейронной сети. Представляет интерес оценка влияния использования полученных аппроксимаций функций активации на работу нейронной сети в целом.

Одним из сценариев использования гомоморфного шифрования является защита уже обученной модели, например, от извлечения данных об обучающей выборке. В этом сценарии сначала обучается стандартная модель, после этого происходит замена нелинейных функций полиномами. Затем к нейронной сети применяется гомоморфное шифрование, после чего она может быть использована для безопасной обработки данных.

В этой связи представляет интерес оценка ошибки результата вычисления выходного значения зашифрованной нейронной сети относительно исходной сети, вызванной заменой функций активации приближениями. В работах [2, 3, 10] такая задача в явном виде не рассматриваются – авторы экспериментально оценивают точность классификации сети на стандартных наборах данных.

В работе [11] авторы предлагают общий подход, основанный на представлении нейронной сети в виде композиции элементарных преобразований, для каждого из которых вычисляется оценка ошибки аппроксимации. Показано, что модуль разности значения выхода исходной сети и ее аппроксимации может быть ограничен величиной $Const \cdot 2^{-\alpha}$, где α – параметр, а $Const$ – некоторая константа. Возникает задача оценки значения этой константы.

Так в работе [14] отмечается, что подход, использованный в [11] и основанный на вычислении композиции преобразований, может приводить к завышению оценки. Это следует из того факта, что если функция использует какую-либо переменную несколько раз, а такое справедливо для нейронных сетей с большим числом слоев, ошибки, вносимые данной переменной, потенциально могут компенсироваться. Для более точной оценки может быть использован описанный в той же монографии аналитический способ, однако он требует вычисления частных производных от функции, описывающей нейронную сеть целиком, что в общем случае может являться трудоемкой задачей. Вместо этого для конкретной нейронной сети мы можем оценить такую ошибку экспериментально, вычислив оценку величины $max_loss(F, F^*)$. Здесь F, F^* – функции, описывающие исходную и приближенную нейронные сети соответственно, а максимум берется по множеству входных аргументов.

Рассмотрим простейшую сверточную нейронную сеть для распознавания изображений цифр от 0 до 9 из базы данных MNIST [15] и приближенный вариант такой сети, для которых вычислим указанное значение max_loss . Каждое изображение является монохромным и представляет собой квадрат со стороной 28 пикселей. Будем считать, что каждый пиксель принимает значение от -200 до 200, что позволяет пользоваться полученными выше оценками.

Нейронная сеть реализуется последовательным применением:

1. Сверточного слоя с ядром свертки 5×5 пикселей, преобразующего матрицу $A = (a_{ij})_{28 \times 28}$ в матрицу $B = (b_{ij})_{24 \times 24}$ с последующим масштабированием значений элементов в указанный выше интервал ($b_{ij} \in [-200, 200]$).

2. Слоя функции активации $ReLU$, которая применяется к каждому элементу полученной матрицы.

3. Слоя субдескрипции, дающего на выходе матрицу $C = (c_{ij})_{12 \times 12}$, $c_{ij} = 1/4(ReLU(b_{2i,2j}) + ReLU(b_{2i+1,2j}) + ReLU(b_{2i,2j+1}) + ReLU(b_{2i+1,2j+1}))$, $c_{ij} \in [-200, 200]$.

4. Слоя регуляризации, исключающего случайным образом 20% нейронов, в результате чего в общей сложности остается 116 элементов,

подлежащих дальнейшей обработке. Без ограничения общности будем считать, что выбираются первые 116 элементов: $D = (d_1, \dots, d_{116})$, $d_{ij} \in [-200, 200]$.

5. Полносвязного слоя из 116 нейронов с функцией активации $ReLU$, $E = (e_1, \dots, e_{116})$. $e_i = ReLU(\sum_{j=1}^{116} \gamma_{ij} d_j)$, $e_i \in [0, 200]$.

6. Финального полносвязного слоя из 10 нейронов с многомерной логистической функцией ($Softmax$): $F = (f_1, \dots, f_{10})$, $f_i = \frac{e^{z_i}}{\sum_{j=1}^{10} e^{z_j}} \in [0, 1]$, где $z_i = \sum_{j=1}^{116} \kappa_{ij} e_j$, при этом значения z_i нормируются: $z_i \in [-1, 1]$.

Эксперимент по оценке возникающей при аппроксимации функций $ReLU$ ошибки рассмотренной нейронной сети в целом заключался в следующем:

1. Случайным образом вырабатывались весовые коэффициенты нейронов.

2. Случайным образом вырабатывалось 10^6 входных матриц A .

3. Для каждой из матриц вычислялось значение $|f_1(A) - f_1(A)^*|$, где f_1 и f_1^* – значение первого выходного нейрона для исходной нейронной сети и ее аппроксимации на одной и той же входной матрице A .

4. Вычислялось значение функции $max_loss(f_1, f_1^*) = \max_A |f_1(A) - f_1(A)^*|$.

Аппроксимация проводилась с помощью полинома четвертой степени, дающего наилучшее приближение (см. Таблицу 1). Функция max_loss не аппроксимировалась, поскольку она не содержит информации о параметрах нейронной сети.

Проведенный эксперимент показал, что $max_loss(f_1, f_1^*) = 0.200293$.

Таким образом, даже для рассмотренной сравнительно простой сети использование аппроксимаций функций активации может приводить к достаточно большим искажениям выходных значений.

6. Заключение. Полученные результаты не позволили в точности получить приближения, опубликованные в широко цитируемой работе [3], или используемые в доступной в сети Интернет программной реализации [10], подготовленной на ее основе. Однако точность получаемых аппроксимаций в целом согласуется с опубликованными ранее результатами. Для функции $ReLU(x)$ удалось получить представление аппроксимирующего полинома, обладающего большей точностью, чем в рассмотренных источниках, в смысле использованной метрики точности приближения ρ . Получен вид аппроксимирующих полиномов для функции логистический сигмоид.

Полученные результаты показывают, что наилучшее полиномиальное приближение из трех рассмотренных функций активации можно построить для функции $ReLU(x)$ с помощью численного метода, что обусловлено ее поведением на числовой прямой, а именно неограниченным возрастанием. Точность полученных с помощью полиномов Чебышева наилучших полиномиальных аппроксимаций функций $\sigma(x)$ и $\tanh(x)$ ниже из-за наличия асимптот при $x \rightarrow \pm\infty$. Таким образом, для снижения ошибок классификации при защите нейронных сетей с помощью алгоритмов гомоморфного шифрования целесообразно использовать сети, состоящие из нейронов с функцией активации $ReLU(x)$.

Для простейшей сверточной нейронной сети, включающей два слоя нелинейных преобразований с функцией активации $ReLU(x)$, приведена экспериментальная оценка отклонения в значении одного выходного нейрона ее приближенного варианта относительно исходного.

В отличие от опубликованных ранее работ нами исследованы различные подходы к построению аппроксимаций функций активации. Результаты работы (полученные приближения или методы построения наилучших приближений для рассмотренных функций) целесообразно использовать при реализации методов конфиденциального машинного обучения.

Литература

1. Pitropakis N., Panaousis E., Giannetsos T., Anastasiadis E., Loukas G. A taxonomy and survey of attacks against machine learning // *Comput. Sci. Rev.* 2019. Vol. 34. pp. 100 – 199.
2. Dowlin N., Gilad-Bachrach R., Laine K., Lauter K., Naehrig M., Wernsing J. CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy // *Proceedings of the 33rd International Conference on Machine Learning (ICML)*. 2016. pp. 201 – 210.
3. Hesamifard, E., Takabi, H., Ghasemi, M. CryptoDL: Deep neural networks over encrypted data // *arXiv preprint:1711.05189*. 2017.
4. Juvekar C., Vaikuntanathan V., and Chandrakasan A. GAZELLE: A low latency framework for secure neural network inference // *27th USENIX Security Symposium*. USENIX Association. 2018. pp. 1651—1669.
5. Brakerski Z., Gentry C., Vaikuntanathan V. (Leveled) Fully Homomorphic Encryption Without Bootstrapping // *ACM Trans. Comput. Theory*. vol. 6. 2014. pp. 13:1–13:36.
6. Cheon J.H., Kim A., Kim M., Song Y. Homomorphic encryption for arithmetic of approximate numbers // *Proceedings of the International Conference on the Theory and Applications of Cryptology and Information Security*. 2017. pp. 409 – 437.
7. Crawford L. H., Gentry C., Halevi S., Platt D., and Shoup V. Doing realwork with FHE: The case of logistic regression // *6th Workshop Encrypted Comput. Appl. Homomorphic Cryptogr. (WAHC)*, New York, USA. 2018. pp. 1–12.
8. Ghodsi Z., Gu T., Garg S. Safetynets: Verifiable execution of deep neural networks on an untrusted cloud // *Advances in Neural Information Processing Systems*. 2017. pp. 4672—4681.

9. Ramy E.A., Jinhyun S., Salman Avestimehr. On Polynomial Approximations for Privacy-Preserving and Verifiable ReLU Networks // arXiv preprint:2011.05530, 2020.
10. Репозиторий проекта CryptoDL URL: <https://github.com/inspire-lab/CryptoDL> (дата обращения: 01.09.2021).
11. Lee J., Lee E., Lee J.-W., Kim Y., Kim Y.-S., No J.-S. Precise Approximation of Convolutional Neural Networks for Homomorphically Encrypted Data // arXiv preprint:2105.10879, 2021.
12. Krogh F.T. Efficient Algorithms for Polynomial Interpolation and Numerical Differentiation, Math. Comput. 1970. vol. 24. no. 109. pp. 185 – 190.
13. Молчанов И.Н. Машинные методы решения прикладных задач. Алгебра, приближение функций. Киев: Наук. думка, 1987. С. 288.
14. Taylor J.R. An introduction to error analysis, University Science Books Mill Valley, California, 1982. P. 344.
15. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998. vol. 86, no. 11. pp. 2278–2324.

Маршалко Григорий Борисович — эксперт, Технический комитет по стандартизации "Криптографическая защита информации". Область научных интересов: защита информации, криптография, биометрическая идентификация. Число научных публикаций – 32. marshalko_gb@tc26.ru; ул. Отрадная, 2В-1, 127273, Москва, Россия; р.т.: +7(495)737-6192.

Труфанова Юлия Анатольевна — эксперт, Технический комитет по стандартизации "Криптографическая защита информации". Область научных интересов: защита информации, биометрическая идентификация, машинное обучение. Число научных публикаций – 1. trufanova_ua@tc26.ru; ул. Отрадная, 2В-1, 127273, Москва, Россия; р.т.: +7(495)737-6192.

G.B. MARSHALCO, J.A. TRUFANOVA
**POLYNOMIAL APPROXIMATIONS FOR SEVERAL NEURAL
NETWORK ACTIVATION FUNCTIONS**

Marshalko G.B., Trufanova J.A. Polynomial Approximations for Several Neural Network Activation Functions.

Abstract. Active deployment of machine learning systems sets a task of their protection against different types of attacks that threaten confidentiality, integrity and accessibility of both processed data and trained models. One of the promising ways for such protection is the development of privacy-preserving machine learning systems, that use homomorphic encryption schemes to protect data and models. However, such schemes can only process polynomial functions, which means that we need to construct polynomial approximations for nonlinear functions used in neural models. The goal of this paper is the construction of precise approximations of several widely used neural network activation functions while limiting the degree of approximation polynomials as well as the evaluation of the impact of the approximation precision on the resulting value of the whole neural network. In contrast to the previous publications, in the current paper we study and compare different ways for polynomial approximation construction, introduce precision metrics, present exact formulas for approximation polynomials as well as exact values of corresponding precisions. We compare our results with the previously published ones. Finally, for a simple convolutional network we experimentally evaluate the impact of the approximation precision on the bias of the output neuron values of the network from the original ones. Our results show that the best approximation for ReLU could be obtained with the numeric method, and for the sigmoid and hyperbolic tangent – with Chebyshev polynomials. At the same time, the best approximation among the three functions could be obtained for ReLU. The results could be used for the construction of polynomial approximations of activation functions in privacy-preserving machine learning systems.

Keywords: activation function, ReLU, tanh, sigmoid, homomorphic encryption, BGV, CKKS, neural network, polynomial approximation, privacy-preserving machine learning.

Marshalko Grigory — Expert, Technical committee for standardization "Cryptography and security mechanisms". Research interests: information security, cryptography, biometric identification. The number of publications – 32. marshalko_gb@tc26.ru; 2B-1, Otradnaya St., 127273, Moscow, Russia; office phone: +7(495)737-6192.

Trufanova Julia — Expert, Technical committee for standardization "Cryptography and security mechanisms". Research interests: information security, biometric identification, machine learning. The number of publications — 1. trufanova_ua@tc26.ru; 2B-1, Otradnaya St., 127273, Moscow, Russia; office phone: +7(495)737-6192.

References

1. Pitropakis N., Panaousis E., Giannetos T., Anastasiadis E., Loukas G. A taxonomy and survey of attacks against machine learning // *Comput. Sci. Rev.* 2019. Vol. 34. pp. 100 – 199.
2. Dowlin N., Gilad-Bachrach R., Laine K., Lauter K., Naehrig M., Wernsing J. *CryptoNets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy // Proceedings of the 33rd International Conference on Machine Learning (ICML).* 2016. pp. 201 – 210.

3. Hesamifard, E., Takabi, H., Ghasemi, M. CryptoDL: Deep neural networks over encrypted data // arXiv preprint:1711.05189, 2017.
4. Juvekar C., Vaikuntanathan V., and Chandrakasan A. GAZELLE: A low latency framework for secure neural network inference // 27th USENIX Security Symposium. USENIX Association, 2018. pp. 1651—1669.
5. Brakerski Z., Gentry C., Vaikuntanathan V. (Leveled) Fully Homomorphic Encryption Without Bootstrapping // ACM Trans. Comput. Theory. vol. 6. 2014. pp. 13:1–13:36.
6. Cheon J.H., Kim A., Kim M., Song Y. Homomorphic encryption for arithmetic of approximate numbers // Proceedings of the International Conference on the Theory and Applications of Cryptology and Information Security. 2017. pp. 409 – 437.
7. Crawford L. H., Gentry C., Halevi S., Platt D., and Shoup V. Doing realwork with FHE: The case of logistic regression // 6th Workshop Encrypted Comput. Appl. Homomorphic Cryptogr. (WAHC), New York, USA. 2018. pp. 1–12.
8. Ghodsi Z., Gu T., Garg S. Safetynets: Verifiable execution of deep neural networks on an untrusted cloud // Advances in Neural Information Processing Systems. 2017. pp. 4672—4681.
9. Ramy E.A., Jinhyun S., Salman Avestimehr. On Polynomial Approximations for Privacy-Preserving and Verifiable ReLU Networks // arXiv preprint:2011.05530, 2020.
10. CryptoDL repository URL: <https://github.com/inspire-lab/CryptoDL> (accessed: 01.09.2021).
11. Lee J., Lee E., Lee J.-W., Kim Y., Kim Y.-S., No J.-S. Precise Approximation of Convolutional Neural Networks for Homomorphically Encrypted Data // arXiv preprint:2105.10879, 2021.
12. Krogh F.T. Efficient Algorithms for Polynomial Interpolation and Numerical Differentiation, Math. Comput. 1970. vol. 24. no. 109. pp. 185 – 190.
13. Molchanov I.N. Mashinnie metody reshenia prikladnikh zadach. Algebra, priblizhenie funkcij [Machine methods of solving applied tasks. Algebra, function approximation]. Kiev: Nauk. dumka. 1987. P. 288. (in Russ.)
14. Taylor J.R. An introduction to error analysis, University Science Books Mill Valley, California, 1982. P. 344.
15. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998. vol. 86, no. 11. pp. 2278–2324.

O. FUENTES, J. SAVAGE, L. CONTRERAS

A SLAM SYSTEM BASED ON HIDDEN MARKOV MODELS

Fuentes O., Savage J., Contreras L. A SLAM System Based on Hidden Markov Models.

Abstract. Methods of simultaneous localization and mapping (SLAM) are a solution for the navigation problem of service robots. We present a graph SLAM system based on Hidden Markov Models (HMM) where the sensor readings are represented with different symbols using a number of clustering techniques; then, the symbols are fused as a single prediction, to improve the accuracy rate, using a Dual HMM. Our system's versatility allows to work with different types of sensors or fusion of sensors, and to implement, either active or passive, graph SLAM. A graph-SLAM approach proposed by the International's Karto Robotics in Cartographer, the nodes represent the pose of the robot and the edges the constraints between them. Nodes are usually defined according to contiguous nodes except when loop closures are detected where constraints for non-contiguous nodes are introduced, which corrects the whole graph. Detecting loop closure is not trivial; in the ROS implementation, scan matching is performed by Sparse Pose Adjustment (SPA). Cartographer uses an occupancy map in order to estimate the position where the map representation is done via Gmapping. The Toyota HSR (Human Support Robot) robot was used to generate the data set in both real and simulated competition environments. In our SLAM representation, we have wheel odometry estimate according to initial position of the robot, a Hokuyo 2D Lidar scan for observations, and a signal control and a world representation is estimated. We tested our system in the kidnapped robot problem by training a representation, improving it online, and, finally, solving the SLAM problem.

Keywords: localization, SLAM, robot navigation, mapping, Hidden Markov Model, sensor fusion, service robot

1. Introduction. Service robots, such as the Toyota HSR [1], are increasingly becoming a part of our everyday life, so the ability to explore, map, and navigate its surroundings is of the utmost importance. SLAM or simultaneous localization and mapping is a solution for this problem.

There are many accepted and well studied methods for solving SLAM, a brief overview of the main paradigms used for solving SLAM is presented on section 2. Depending on the application, one approach might be better suited than other. The sensors information available is also an important factor to decide which approach to use, e.g. Wheel information is an efficient way of estimating small changes in position, however a drone would not have this valuable information. Taking into consideration these differences may favour a SLAM method or type of sensor in a specific environment or even a specific region in an environment, e.g., a dark corner may be a terrible place to use image-based methods, on the other hand, a colored flat wall would render little information to a LIDAR-based one.

We propose a graph SLAM system based on HMM's (Hidden Markov Models). Our method can take advantage of several types of sensors measure-

ments (or sensor fusion) while estimating a graph topological representation. Dual-HMM allows us to use two different quantizers-simultaneously, effectively fusing data at a symbol level. Furthermore, new nodes can be added without modifying the graph (modularity). They can be grown or altered without training an entirely new model.

In summary, our main contributions are:

- A graph SLAM system based on Hidden Markov Models.
- A modular system capable of using a wide variety of sensors and features.
- An autonomous training method.
- A navigation method capable of obstacle avoidance.
- A robust localization method.

The remaining of the paper is divided as follows. In Section 2 we present a summary of the SLAM problem and the HMM-based approaches. Then, in Section 3 we introduce our probabilistic approach to the graph SLAM problem and in Section 5 we describe the experimental results. Finally, the main findings are discussed in Section 6.

2. Related Work. The three main SLAM paradigms are Kalman Filter, Particle Filter and Pose graph based implementations, and all of these can be found in the most commonly used open-source libraries.

2.1. EKF SLAM. The Extended Kalman Filter SLAM [2, 3] is one of the most accepted SLAM solutions; it consists of three basic operations:

2.1.1. Robot Movement. The agent moves increasing its position uncertainty due to odometry errors.

$$S_t \leftarrow f(s_{t-1}, u, n),$$

where:

- f - motion model,
- S_t - state of the robot at time t ,
- u - control signal,
- n - noise.

2.1.2. Discovery. The agent finds new interesting landmarks, which need to be referenced. The position uncertainty and sensor error readings are modeled using an inverse observation model i.e. where a landmark is in the map, given the scene seen by the robot.

$$L_i = g(S_t, \vec{O}_t, y_i),$$

where:

- g , Direct Observation model,
- L_i - i -th landmark,
- S_t - robot state at time t ,
- y_i - measure of i -th landmark.

2.1.3. Re-Discovery. The agent finds a previously mapped landmark and re-estimates both its position and landmark position.

The extended Kalman filter has a “stage” for each of the above operations, making it a useful estimator for propagating the uncertainty related to the three mentioned actions.

$$y_i = h(S_t, \vec{O}_t, L_i),$$

where:

- h - Indirect Observation model.

The map representation itself is a matrix that stacks vectors of all mapped landmarks on any given robot state

$$map = \begin{bmatrix} S \\ L_1 \\ \cdot \\ \cdot \\ L_I \end{bmatrix},$$

where:

- S - robot state.

2.2. Particle Filtering. One of the most important characteristics of this SLAM approach is the building of an occupancy grid map. This map is later used to achieve localization with solutions similar to de EKF filtering, hence the name Filtering in Particle Filtering. Localization is achieved with Adaptive Monte Carlo Localization AMCL, a member of Markov localization algorithms. Murphy [4] et al. proposed Rao-Blackwellized Particle Filters as a SLAM solution; the key idea of the Rao-Blackwellized particle filter for SLAM is to estimate the joint posterior $P(\vec{x}_t, m \mid \vec{z}_t, \vec{u}_{t-1})$. The Rao-Blackwellized particle filter for SLAM makes use of the following factorization:

$$P(x_{1:t}, m \mid z_{1:t}, u_{1:t-1}) = P(m \mid x_{1:t}, z_{1:t})P(x_{1:t} \mid z_{1:t}, u_{1:t-1}),$$

where x is the pose of the robot, m is the map, Z_t are the observations at time t , and U_t is the control signal.

Importance Sampling Filters (such as Sequential Importance Resampling or SIR) are employed to estimate the posterior $P(x_{1:t} \mid z_{1:t}, u_{1:t-1})$, as mentioned by Grisetti et al. in [4], where "each particle represents a potential trajectory of the robot. Furthermore, an individual map is associated with each sample. The maps are built from the observations, and the trajectory represented by the corresponding particle".

2.3. Graph SLAM. According to Grisetti et al., "one intuitive way of formulating SLAM is to use a graph whose nodes correspond to the poses of the robot at different points in time and whose edges represent constraints between the poses" [6].

From a probabilistic point of view, SLAM can be represented with a sequence of random variables as X (robot pose), M (map features), Y (sensor readings), and U (robot motion). To solve the SLAM problem, we simply use the maximum a posteriori probability:

$$P(X_t, M \mid Y_t, U_t, x_0),$$

for each time step t .

There has been a lot of work on solving this estimation problem [6] [7]. However, special attention has been given to [8], where the authors apply variable elimination techniques to reduce the dimensionality of the optimization problem. In the work presented in [9], a solution to the active SLAM problem is proposed "in scenarios in which some prior information about the environment is available in the form of a topo-metric graph".

2.4. ROS SLAM. As mentioned before, many state-of-the-art solutions can be found as open-source libraries in popular frameworks such as the Robot Operating System (ROS). Here, we present some of the most popular implementations.

2.4.1. Gmapping. Gmapping [10] is based on Rao-Blackwelized Particle Filters proposed by Grisetti et al. in [11]. Particle Filters are a known application of Bayesian Filters in which a large number of importance weighted particles represent the a-posteriori probability; a probabilistic occupancy grid is used as a map representation, and AMCL, Adaptive Monte Carlo Localiza-

tion, is used for localization in this map – AMCL is a member of the Markov localization algorithms family.

2.4.2. HECTOR SLAM. HECTOR SLAM does not use odometry information for the localization; it rather uses high update rates and low distance measurements to estimate the robot movement. A version of ICP is used to estimate the pose between samples and maintain the robot pose estimate.

2.4.3. CARTOGRAPHER. A graph-SLAM approach proposed by the International’s Karto Robotics in [12] – in Cartographer, the nodes represent the pose of the robot and the edges the constraints between them. Nodes are usually defined according to contiguous nodes except when loop closures are detected where constraints for non-contiguous nodes are introduced, which corrects the whole graph. Detecting loop closure is not trivial; in the ROS implementation, scan matching is performed by Sparse Pose Adjustment (SPA) [13]. Cartographer uses an occupancy map in order to estimate the position where the map representation is done via Gmapping.

As we can see from this brief review, some methods are better suited for specific tasks or robot architectures. Our research is based on the idea that each method has its strengths, and we propose a method that, focusing on the versatility of application, tries to take some of the advantages inherent to these methods.

3. HMM-based graph-SLAM. We propose a versatile graph-slam system based on Hidden Markov Models; the goal is to estimate a topographic graph given noisy sensor measurements and pose estimates. In this section we will briefly describe the core concepts in our implementation.

3.1. Hidden Markov Models. A Hidden Markov Model (HMM) is a two random variable stochastic process in which only one of the random variables is directly observable. In its discrete version, and provided the Markov property [14] is fulfilled, the system dynamics is fully defined by a transition matrix A , an emission matrix B , and, optionally, an initial conditions vector $\vec{\pi}$.

$$a_{ij} = P(S_t = s_j \mid S_{t-1} = s_i),$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ a_{31} & a_{32} & \dots & a_{3N} \\ \cdot & \cdot & \cdot & \cdot \\ a_{N1} & a_{N1} & \dots & a_{NN} \end{bmatrix},$$

$$b_j(k) = P(V_k | q_t = S_j),$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1M} \\ b_{21} & b_{22} & \dots & b_{2M} \\ b_{31} & b_{32} & \dots & b_{3M} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_{N1} & b_{N1} & \dots & b_{NM} \end{bmatrix},$$

$$\lambda = [A, B, \vec{\pi}].$$

We propose the hidden random variable to represent the pose of the robot, and the observable variable will be a discrete representation (quantization) of the sensor measurements. The dynamic programming based algorithms' family related to HMM's [15] provide efficient estimations to probabilities related to the localization problem.

Our approach on Markov localization uses a discrete code-book to represent the sensor's readings. The symbolic representation of the readings allows for any sensor or fusion of sensors to be used as observations. Localization is achieved, as any other Markov localization approach, by estimating the probability of being on state $S_t = s_n$ given a last known state $S_{t-1} = s_n$, a sensor(s) observation symbol $O_t = V_k$, and a control signal $U(t)$. The 2D pose state space is defined by the tuple $\langle x, y, \theta \rangle$.

For each time step t , random variable S can take the value of a discrete set of N states; similarly, the observations of each time step will be a member of the discrete code-book observation set of K symbols V . Control signal $U(t)$ is assumed constant, and only an on-off signal for the potential fields autonomous navigation system is necessary, as follows:

$$S_{t-1} = s_n, O_t = V_k, U(t) = u.$$

An action set is used for the active slam pose belief exploration. The planning path is then obtained using a search algorithm like Dijkstra or A*. Pose traces are proposed, so given a state belief and an action the most likely future state is estimated.

The algorithms typically associated with HMM's can efficiently estimate this probability using dynamic programming. Even though a number of states

in the models are reported as fixed, it was done to have uniformity between models, since models are modular and can easily be connected to form a global model, similar to sub-maps in Cartographer. The balance between computer costs and accuracy can be fine-tuned with the number of states, and the observations code-book size.

The resulting representation of the environment is a symbolic pose graph, where each node or state is the centroid of the free region surrounding it; orientation is also accounted for with a scaling factor. The HMM model is completed by recording all the observations in the code-book sensed while the agent is at any given state. As the agent acquires more information by exploration, new states are formed, or the existing ones are moved to account for new data, eventually including all the explored areas. The number of total states is a hyperparameter of the model, and it can be seen as a scale factor to be used. The more states in which free space is divided, the more accurate the correction is at the expense of more computational expenses.

One of the main advantages of the method is adaptability, since it can use any sensor or fusion of sensors. As briefly mentioned above, observations are not used directly. A symbol, part of a discrete data set, is used to represent sensor readings. This quantization of readings allows the method to be used with various sensors, different sensor coding, or even sensor fusion. In HMM terms, we have various emission matrices for each transition matrix. We perform estimation on various models, each related to each sensor, or fusion of sensors. The model likelihood given a set of observations is easy to obtain with HMM algorithms like Forward Algorithm, (eq. 2) and its Backward counterpart,(eq. 3); and it is used as a metric to decide which model to use on a specific area. A lidar will gather more information from a dark corner, while a camera would do better on a long hallway with distinctive visual landmarks. Computational cost is in the order of N^2T (eq. 4). However, the number of states rapidly increases the computational cost, "the curse of dimensionality", where the direct calculations without dynamic programming algorithms are in the order of N^T (eq. 1).

$$P(\vec{O} | \lambda) = \sum_{j=1}^S \left(\pi_{s1} \prod_{t=1}^{i=T} a_{s_t, s_{t+1}} b_{s_t - s_{t+1}}(o_i) \right), \quad (1)$$

$$\alpha_{t+1}(i) = \sum_{j=1}^N (\alpha_t(i) a_{i,j}) b_j(o_{t+1}), \quad (2)$$

$$\beta_t(i) = \sum_{j=1}^N (a_{i,j} b_j(o_{t+1}) \beta_{t+1}(j)), \quad (3)$$

$$P(\vec{O} | \lambda) = \sum_{i=1}^N \alpha_i(i) \beta_i(i). \quad (4)$$

Finally, localization is performed using the Viterbi algorithm; this algorithm estimates the most probable sequence of states given a sequence of observation symbols. The model with maximum likelihood is obtained with the Forward Algorithm. The most likely model given the readings is used in the Viterbi algorithm, with $\max [P(\vec{S} | \lambda, \vec{O})]$ being the discrete representation of the space in which each state is the closest centroid to the robot state at the time of sensor capture. The size of the free region represented by each state is variable, but it is kept to a small enough area to make wheel odometry reliable. The transitions between states happen on well-defined regions normally distributed around a transition point. The most frequent transition point between states is used as a "reset" wheel odometry according to that point. Similarly to cartographers' sub-maps, wheel odometry is reliable in small regions.

4. Implementation. We have presented all the components of the system we propose in a very general way. Now, we will show some interesting implementations using those components. It is important to mention that a map is not necessary for our method but, if one is available or needed, given that our method deals with the same probabilities as the particle filter used in AMCL [16], it is possible to use Rao-Blackwellization. Furthermore, as our training methods create a decent grid estimation, whether using a map or not is greatly dependant on the application and error scale needed.

In our implementation, a global reference frame is created. However, global mapping based on iterative closest point (ICP) methods is also a possibility we explored in some regions, especially when a high.

In our SLAM representation, we have wheel odometry estimate according to initial position of the robot, a Hokuyo 2D Lidar scan for observations, and a signal control and a world representation is estimated. The HMM Model is trained as follows.

4.1. Training. Short exploration runs are conducted, gradually restarting to a known value, like an entrance or a recharging station; this episodic learning approach makes the wheel odometry reliable. The duration of the episodes is such that wheel odometry is reliable. The control signal in exploration mode is an on/off signal enabling a reactive potential fields behavior [17]. Some additional constraints are added to the behavior, as mentioned in our previous work [18], like an exploratory turn every given time sample, or an artificial attraction towards doors.

Model training is done with a labeled training set. This training set is formed by a vector \vec{O}_t that contains all the 720 laser readings from a laser sensor and an odometry-based pose vector or hidden state S_k .

Baum-Welch [15] is the most commonly used method for the HMM model estimation. It is an elegant approach that calculates the optimal model parameters given the observations; it is also an Expectation-Maximization (EM) algorithm [19]. The optimization is done by maximizing the Likelihood of the Model given some readings. The most common training method is used offline and will serve as an initial estimation for our method. We maximize the likelihood of the model given some observations (training set). An auxiliary variable ξ is defined.

$$\begin{aligned}\xi_t(i, j) &= P(s_t = i, s_{t+1} = j | o_t, \lambda), \\ \xi_t(i, j) &= \frac{\alpha_t(i) a_{ij} b_j(o_{T+1}) \beta_{t+1}(j)}{P(o | \lambda)} \\ &= \frac{\alpha_t(i) a_{ij} b_j(o_{T+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(o_{T+1}) \beta_{t+1}(j)}, \\ \gamma_t(i) &= \sum_{j=1}^N \xi_t(i, j), \\ \sum_{t=1}^{T-1} \gamma_t(i) &= S_i^T,\end{aligned}$$

where S_i^T is the number of transitions from S_i over T time steps.

$$\sum_{t=1}^{T-1} \xi_t(i, j) = S_{ij}^T,$$

where S_{ij}^T is the number of transitions from S_i to S_j over T time steps.

$$\begin{aligned}\hat{a}_{i,j} &= \frac{S_{ij}^T}{S_i^T} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(j)}, \\ \hat{b}_i(k) &= \frac{\sum_{t=1}^T 1_{o_t=v_k} \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}.\end{aligned}$$

This variable can be easily represented with the Backward and Forward variables, as follows [15] uses $\sum_{t=1}^T 1_{o_t=v_k}$ to represent number of times V_k is seen in state j .

We also use a gradient descent implementation of the Baum-Welch algorithm [20] to keep optimizing the model online. This can introduce subtle changes when the environment changes or when a new area is mapped and added to the global model. A mini-batch approach can also be used to re-estimate the model every time a batch buffer is filled. Each new reading contributes to an infinitesimal change in its respective row and column in the transition matrix. Note from eqs. 5 and 6 that changes only happen on the row and column related to the reading. The least frequent poses will be absorbed by the more frequent (desired) readings. This online process allows aligning the new readings and reinforcing the optimal values given a space and obstacles configuration into the matrix.

$$\frac{\partial P}{\partial a_{ij}} = \sum_{t=1}^T \alpha_t(i) b_j(O_{t+1}) \beta_{t+1}(j), \quad (5)$$

$$\frac{\partial P}{\partial b_j(O_t)} = \begin{cases} \delta_{j1} \beta_1(j) & t = 1 \\ \sum_{i=1}^N \alpha_{t-1}(i) b_{ij}(O_{t+1}) \beta_t(j) & t \neq 1 \end{cases} \cdot \quad (6)$$

Due to the low complexity for estimating a model and, later on, navigating the model to obtain information, it is possible to use several models simultaneously; this allows us to add new nodes (i.e. states) to the model without the need to re-estimate the whole model.

In case autonomous mapping is not required, and human interaction is used to "explore" the scene, the model will benefit from different strategies, and active policy estimate can be obtained. Again, versatility is an important characteristic of our method. Regardless of how the training set is obtained and the model estimated, that is, either offline, online, batches, etc., an HMM is obtained with the Baum-Welch algorithm. Figure 1 shows a topological node graph obtained from the transition matrix of the proposed HMM after online navigation. Results found in Section 5 and Figure 10 and Figure 11 show this process in a standard competition arena where an apartment is explored, and an initial graph is estimated; the green arrows represent the state $\langle x, y, \theta \rangle$. The

background image is just a reference for illustrative purposes since the map is not being used.

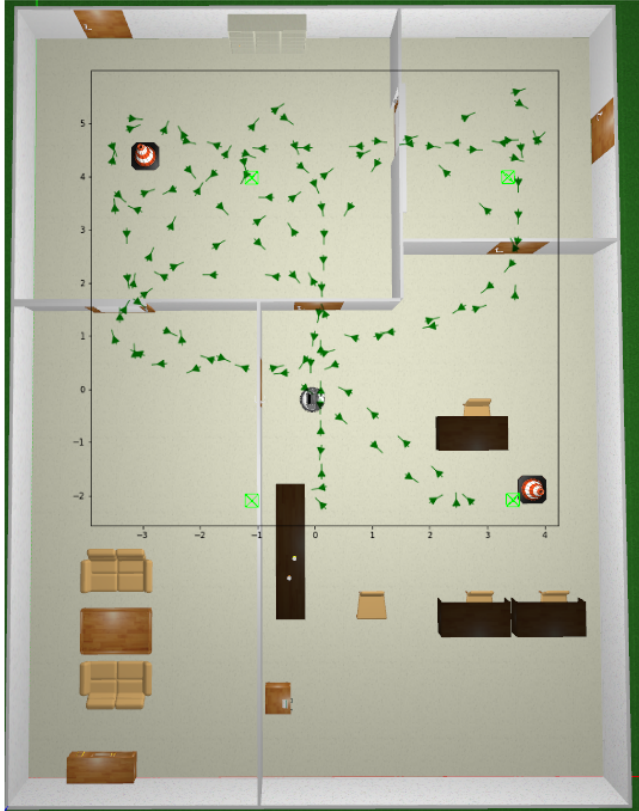


Fig. 1. Topographical graph representation, obtained from the transition matrix of the HMM trained. Graph is created with the states corpus (2D poses) shown as green arrows

4.2. Observations. Observations used for the estimation of the HMM are quantized (and finite) in nature, so the method used for the quantization or the kind of the sensor itself is not important as long as we use a finite code-book to represent them. We explored various observation alphabets for

the code-book where Lidar readings were treated using different clustering techniques, namely K-means and affinity propagation (Figure 2).

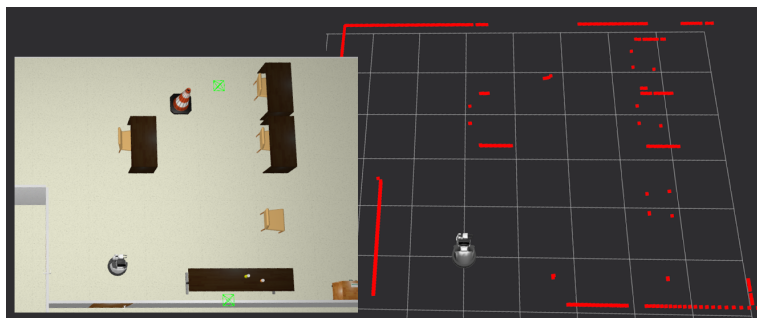


Fig. 2. Apartment scene with its corresponding Hokuyo Lidar reading

Some experiments were also done by fusing LIDAR and 2D images features obtained with Resnet [21] architecture. The final pose estimator efficiency greatly depends on good readings. "A good reading is that which can be re-observed, and one easily differentiates from other reads" [22].

4.2.1. Lidar. K-means (or mini K-means in this case) works well when a large amount of data is used. The low occurrence of outliers takes them out of the estimation; however, highly repeated values will "skew" the means resulting in some information loss. A different algorithm is proposed (Affinity Propagation) [23] to quantize the laser readings based on their similarity, rather than their frequency of appearance. A dual HMM is introduced, which estimates independently using code-books from each clustering algorithm, and trusting only matching estimates obtained independently by each model. A block diagram of the proposed DUAL HMM architecture can be seen in Figure 3. The statistical characteristics of most readings are quite similar. This similarity enables the usage of affinity propagation clustering techniques to generate more heterogeneous clusters; however, the drawback is the lack of scalability when there are many samples. On the other hand, K-means works well with large data sets and being a Euclidean distance approach, it can be used online (with the previous centroids).

In sum, we use measurement symbols as anchors of information and depending on the sensor's nature, on the environment conditions, and even on the route, one kind of symbol generation might result in a better performance.

A robust SLAM system should benefit from all symbol representations; Sensor Fusion [24] is used in this way, yet we propose a much simpler approach:

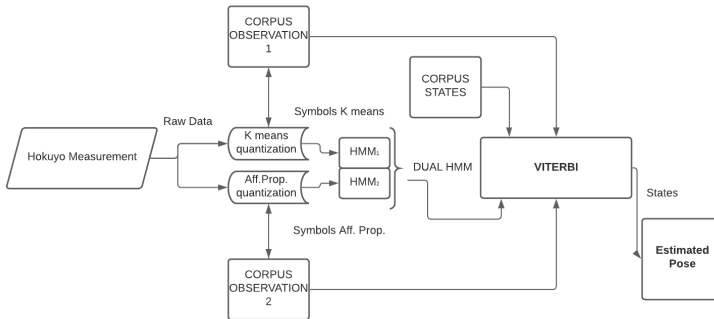


Fig. 3. Block diagram of the proposed DUAL-HMM system

Dual-HMM. In our experiments, we use two different features extracted from the same observations vector, but this applies to any kind of observation.

Figure 4 shows a common scene of a service robot, and on the right side is an example of a typical lidar measurement in green; the centroid obtained with K means in orange and the Affinity Propagation exemplar in green.

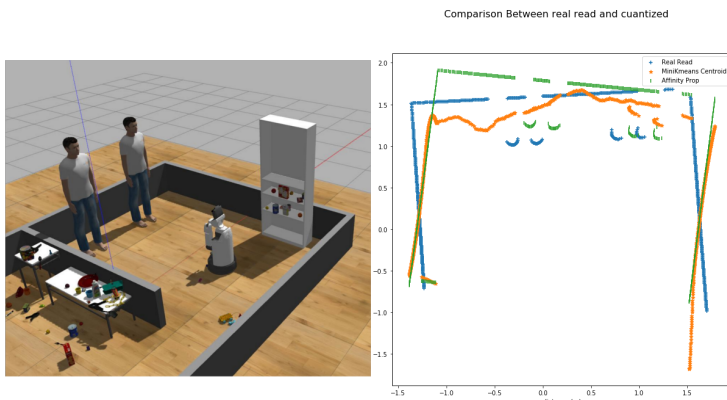


Fig. 4. Example of a Hokuyo typical reading (blue) with its K-means centroid (orange), and affinity propagation exemplar (green)

The final pose estimator's efficiency greatly depends on good readings "a good reading is that which can be re-observed, and one easily differentiates from other readings" [22]. K-means (or mini K-mean, in this case) works well when a large amount of data is used since the low occurrence of outliers takes

them out of the estimation; however, highly repeated values will skew the means. Affinity Propagation [23] is another clustering algorithm, and it was used to quantize the Hokuyo readings based on their similarity rather than their frequency of appearance. Figure 5 shows K-means information loss in the most observed symbol. In comparison, Figure 6 shows the symbols assigned by the more computationally expensive Affinity Propagation to a series of new readings.

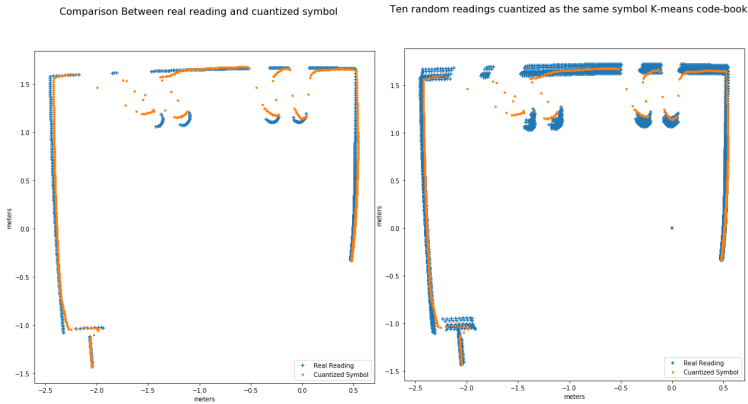


Fig. 5. Most common read and 10 samples clustered in the same K-means alphabet symbol

In our approach, a dual HMM makes independent estimates using symbols from each clustering algorithm. Although there are other methods to achieve sensor fusion, in our proposal, we use a simple symbol level data fusion.

4.2.2. Resnet feature extraction. Another observations alphabet was quantized by getting an observation vector \vec{O}_t from the last layer of a Resnet 50 CNN (Convolutional Neural Network). Such vectors were quantized into an alphabet and their respective emission matrix estimated. All models share the same transition matrix. Results reported in section 5 show models using the fusion of a Resnet symbol with a lidar K-means symbol.

4.3. Localization. Once a model, or several models, are found, it is possible to estimate the robot's pose with a set of past observation symbols $\vec{O} = V_t, V_{t-1}, \dots, V_{t-M}$. The most direct approach is to the Viterbi algorithm [25] that takes a sequence of quantized observations \vec{O} , an HMM model λ , and an initial conditions vector $\vec{\pi}$, and yields the most probable state sequence traversed by the system given the observations and HMM model.

$$P_1 = \pi_i \cdot b_i(o_1),$$

$$P_t(j) = \max_{1 < i < N} [P_{t-1}(i) \cdot a_{ij}] \cdot b_j(o_T),$$

$$S_t(j) = \arg \max_{1 < i < N} [P_{t-1}(i) \cdot a_{ij}],$$

$$P(\vec{S} | \vec{O}, \lambda).$$

The length M of this observations vector is called the Viterbi buffer, and it is a hyperparameter that should be trained to optimize the models. Forward and Backward algorithms are used to find the most common state sequence and the probabilities of being in a specific state given the observations.

4.4. Pose Correction. Wheel odometry is an easy way to estimate a pose with the information from the wheels and the control signals. This estimation is reliable in the short term because a small error – due to slips and control noise – is accumulated over time, i.e. a reference point must be maintained. The accumulated error makes the estimation unreliable in the long run; this error is normally distributed [26] so, it is possible to characterize a particular floor-wheel interaction odometry error.

In our proposal, wheel odometry is corrected by resetting the reference point where the correction is made when a state transition is detected via the

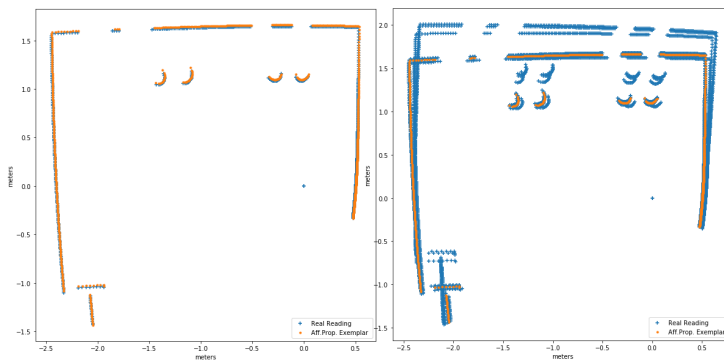


Fig. 6. Most common read and 10 samples clustered in the same affinity propagation alphabet symbol

Viterbi algorithm. It is easy to observe from the training set that the transitions from one state to another usually happen approximately in the same pose value, i.e. transitions happen in a normally distributed point with the same distribution as the wheel odometry error; this value is used as the correction value on trusted transitions. Wheel odometry can be trusted inside the small region related to a specific state or centroid or time interval.

Figure 7 shows in green the position estimate of the Wheel odometry in a run traversing through two states; black points show the real pose of the robot. Two independent HMM's correct the position according to their own estimates – it can be seen in the upper part both models estimate the same pose, and correction is the same.

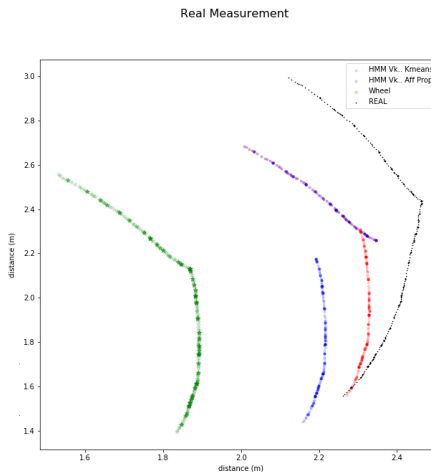


Fig. 7. Real pose in black and three different pose estimation methods. Wheel odometry in green, a single HMM (K-means) in blue and single HMM (aff.prop.) in blue

4.5. Navigation. After the training process, the possible states of the system are represented with the pose centroids corpus; the topological map used by the Dijkstra algorithm is a version of the HMM's transition matrix. We propose different kinds of nodes; some of them are bidirectional, others one-directional, some others only landmarks used to re-estimate the position. This is called *labeling* in some literature relating to HMM and must not be confused with the labeling of the training set. Again, the graph nodes (or pose centroids) are re-estimated online with each reading, and they tend to align with the optimal route as the robot navigates by being attracted to a single centroid at a time (virtual attractor). Dijkstra route will be then a path containing the

sequence of centroids to visit to get from a current position to a goal. Potential fields reactive behavior is used to sequentially visit these virtual attractors while avoiding unmapped or dynamic obstacles. Figure 8 shows a sequence of images of the robot traversing a path made of such attractors – the control signal is simply an on/off signal, and the robot is either trying to get to the next attractor or it is static on a fixed position.

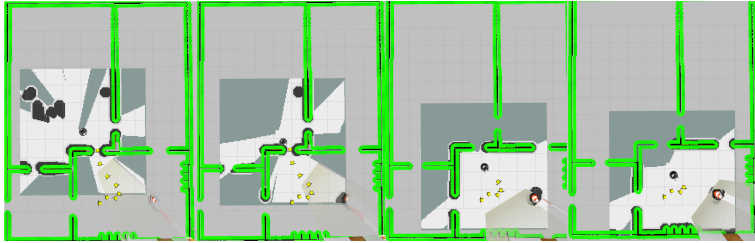


Fig. 8. Different moments of reactive navigation through a route formed by virtual attractors. Route is produced after planning using Dijkstra to navigate the joint transition matrix of the Dual-HMM

5. Experiments. We present results using a Gazebo simulated home environment, a typical indoor setting for service robots [27]; even though a map is provided with the simulated environments, we do not use it in our SLAM tests, except for Figure clarity.

5.1. Benchmark. We use a benchmarking method proposed in [32] to compare our models. Dieter et al. propose "a metric for measuring the performance of a SLAM algorithm not by comparing the map itself but by considering the poses of the robot during data acquisition". As mentioned, a map might be useful for human interaction, and feature comparison, but there are many other representations that can work as a "map" and, although our method could be used to build an occupancy grid with gmapping, the map itself is not required for our SLAM method, and is only presented for clarity.

The proposed metric also allows to compare methods that use a map with others that do not since it only relies on estimates of the trajectory of the robot given by a set of poses where the observations are taken.

5.2. Training. First, the robot roams the environment reactively and, at the same time, registers wheel odometry and laser measurements at every time step. In order to make wheel odometry reliable, episodes are run among known fixed locations like doors or furniture, or in short periods of time – this serves as ground truth comparisons for the localization task and odometry correction metric.

Once enough training data is obtained (around 15k samples in our experiments), two HMM's are calculated, both HMM's share the same transition matrix and, as a consequence, the same topological representation; each HMM, however, uses a different code-book to represent the observations (K-Means and affinity propagation, respectively). Each code-book favors different world areas, so dual representation or sensor fusion can be applied; however, we opted for a dual HMM. The dual approach ensures that only correct estimates are used for correction; as a consequence, wheel odometry is corrected only when both HMM's yield the same estimation given the observations.

The pose can come either from regular uncorrected wheel odometry or, ideally, corrected odometry. In the simulations, Gaussian noise with the parameters obtained from the real robot was added to the ideal training set as a data augmentation pre-processing technique.

Additionally, to test the results of the different models we later perform around 100 sample runs in the arena with no initial information and random routes.

In Figure 9, an example run is shown. Pink points represent older real poses (still unreliable since no initial conditions are used for estimation) and yellow points are poses with a big enough buffer to be reliable. The Figure shows an example of a correct estimation, this problem is commonly known as Kidnapped Robot [30], meaning no initial conditions information is used for the estimates; however, once a correct estimation is made initial, state probabilities are available for further estimates, greatly improving accuracy. Green dots represent the poses centroids, i.e. the corpus of the HMM hidden variable X , and blue and red dots represent real and estimated hidden states, respectively. Wheel odometry correction is not applied at this stage, and accuracy is tested with the quantized states, not the real odometry, i.e. correct estimations on the HMM's with no initial conditions $\vec{\pi}$.

5.3. DUAL-HMM. The simple HMM using K-means Lidar alphabet was enough for the WRS tasks but, in order to test various data fusion schemes, a dual HMM is trained with the same training set – dual HMM uses both Lidar alphabets where the K-means and the Aff. Prop. symbols create each an emission matrix for the same transition matrix. Finally, a fused data model is proposed where the alphabet code-book is the fusion of 2D Lidar readings with the normalized features obtained from the Resnet CNN. Even though some information can be obtained this way, it is highly correlated with the random starting point because, by nature, there are regions with better features than others and, even using initial information, no significant improvement to the model is found.

In our experimental setup, the considerable computer expenses related to CNN's do not make it worth it to solve the robot relocalization problem, since accuracy does not improve significantly. This fusion may favour a different application, robot configuration or environment, and, since we show data fusion using different code-books, it is included in this work as a possible configuration for visual SLAM.

5.4. Modularity and Online Improvement. To show modularity, we use a multiple room environment. New graphs (or states) can be added after an initial HMM was found, or even a completely new HMM for each room; there is no need to retrain the whole model, as our previous research suggests [18]. Figure 10 and Figure 11 show the centroids aligning closer once enough online training had taken place. Online training yields a more structured and organized topological representation of the explored environment (shown in Figure 11), i.e. the topological representation keeps improving as the robot successfully keeps navigating it thanks to on-line Baum Welch algorithm. Figure 10 and Figure 11 show the topological graph representation before and after online improvement.

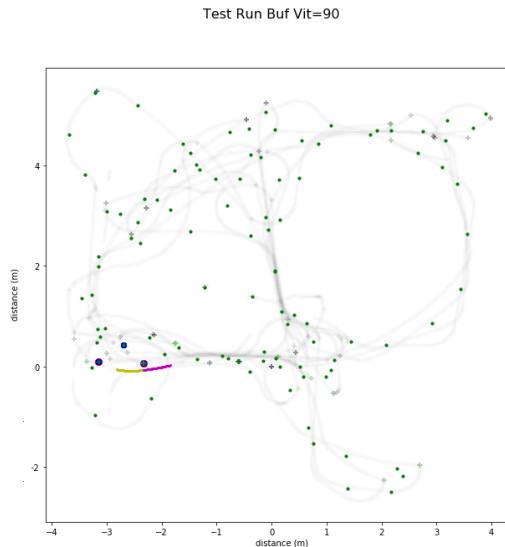


Fig. 9. An example run with a correct estimate and no initial conditions information traversing three states. State corpus in green, estimates on red and blue, beginning of the test run in pink

5.5. Pose Correction. A different virtual environment was used to validate wheel odometry correction. The reality gap is broken since the training data obtained from the model was used to navigate the real world arena, both used in the World Robot Summit 2018 and 2020 (WRS) competitions. Figure 12 and Figure 13 show the arena and the topological graph found.

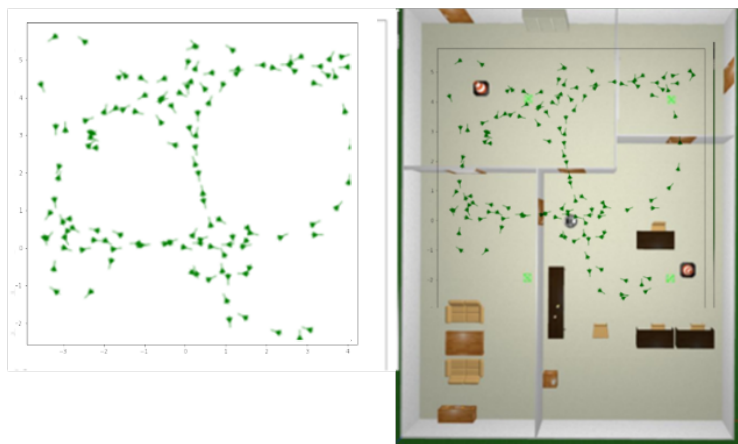


Fig. 10. Initial topological representation of the environment as obtained from off-line training

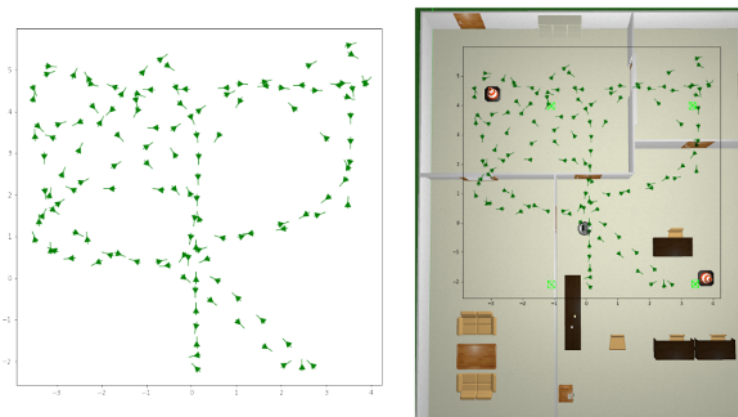


Fig. 11. A more structured and organized topological representation of the explored environment, thanks to on-line training



Fig. 12. World Robot Summit arena with the proposed topological graph representation

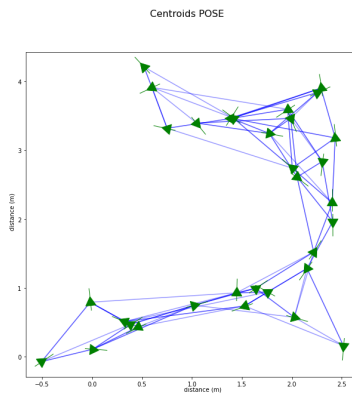


Fig. 13. WRS arena HMM representation Topological Map graph

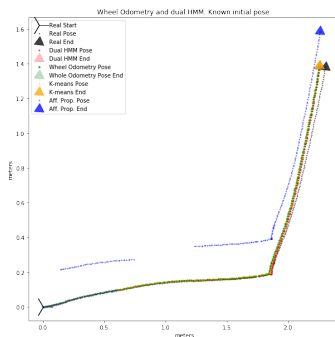


Fig. 14. Comparison with known initial location of Wheel Odometry estimate with known initial conditions in green , Dual-HMM pose estimate in red. Each of the dual single components observations K means in orange and Affinity Propagation in blue. Real trajectory shown in black

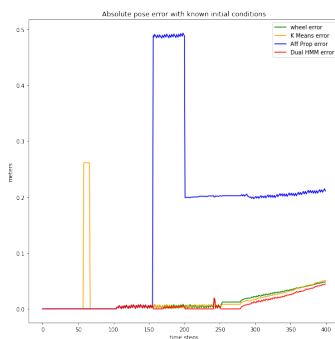


Fig. 15. Absolute pose error of Figure 14

Again, the robot was trained by acquiring the training set on a potential fields autonomous exploration run. Since the dimensions of the arena made wheel odometry reliable enough, the whole arena was modeled in the same HMM transition matrix (Figure 13).

The test runs shown do not use initial state information (kidnapped robot); however, this initial information is available after a good estimate, and could be used for maintaining the agent's pose estimate.

Such a case is shown in Figure 14 where a comparison between the real trajectory and the components (K-means in orange and Affinity Propagation in

blue) of dual HMM, as well as dual HMM in red; the real trajectory is displayed in black.

In case a good enough current initial position is known, wheel odometry is a very good way to estimate pose; once error starts growing, it can be seen that Dual HMM keeps pose estimates better aligned on the long run. It is important to note that HMM has a small quantization error on the first estimation because there is no initial information, and as such no initial transition, just the centroid of the hidden state is available in the first kidnapped robot estimate. A much longer run is shown in Figure 16 to further illustrate this idea. Figure 15 and Figure 17 report the absolute error in pose estimate over time.

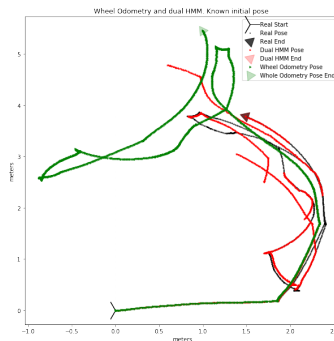


Fig. 16. Same initial conditions than Figure 14 but a much longer run. Wheel odometry in green clearly diverges while Dual-HMM in red remains operational through the whole run

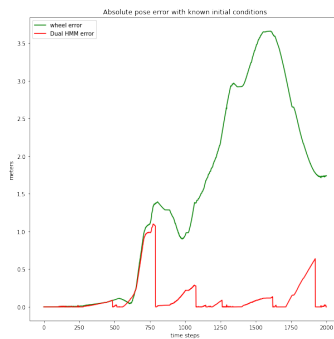


Fig. 17. Absolute pose error of a long run

5.5.1. Kidnapped robot. The wheel odometry has no initial information, and no way to re-estimate its real position, so the error keeps growing as time goes on, unlike the dual HMM odometry; as the agent continues to move it keeps traversing reliable states and correcting its pose. Results shown in Figure 18 and Figure 19 suggest that the kidnapped robot problem was solved once enough time to traverse a trusted transition had elapsed; however, if wheel odometry pose has initial conditions, it is a better estimate for the first few time steps.

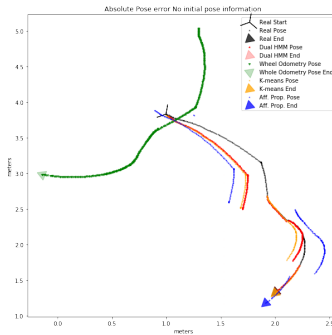


Fig. 18. In case wheel odometry has no initial information

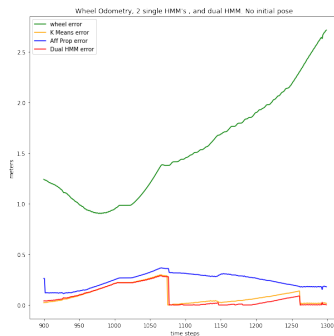


Fig. 19. Absolute pose error of run with unknown initial conditions

5.5.2. Catastrophic Event. An example run before such an event is shown in Figure 20. It can be seen in Figure 21 that the wheel odometry error is smaller than the quantization error. Then, in Figure 22, the beginning of a catastrophic event that greatly increases wheel odometry error is shown. Such

error is can be seen in Figure 23; however, dual HMM SLAM can re-estimate its pose after traversing a reliable transition, since the system resets its odometry to a known transition value, remaining operational on the long run as shown in Figure 24 and Figure 25.

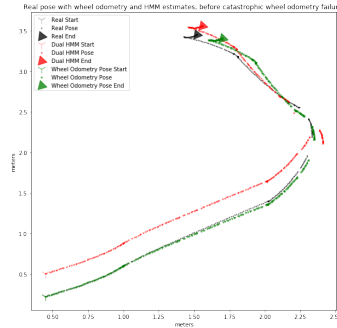


Fig. 20. Wheel Odometry and Dual-HMM error. Event start. In black the real pose sequence, green is wheel odometry with initial conditions, in red the Dual-HMM, which has a quantization error at the beginning of the run before a catastrophic event

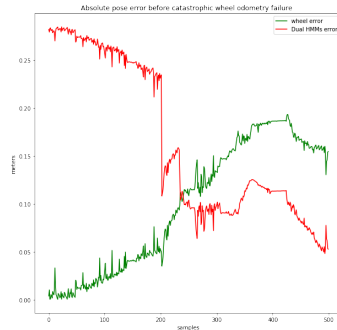


Fig. 21. Absolute pose error of wheel odometry and Dual-HMM. Both errors are relatively small, it is interesting to see how after a few time steps Dual HMM error is similar

6. Conclusions and Future Work. We proposed a SLAM method based on Hidden Markov Models which generated a modular graph representation of the environment; once the model was calculated, the kidnapped robot problem was solved in a number of environments using 2D sensor readings as observations.

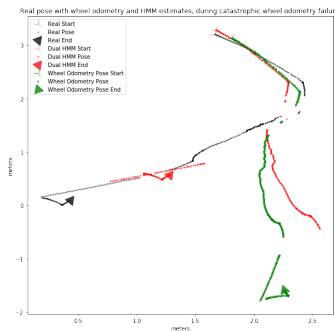


Fig. 22. Wheel Odometry in green and Dual-HMM in red during Event

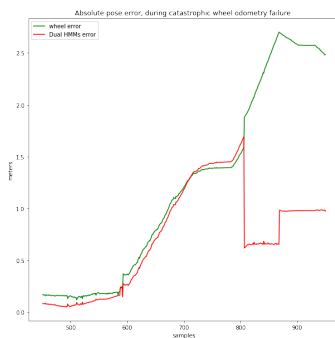


Fig. 23. Wheel Odometry and Dual-HMM error during Event

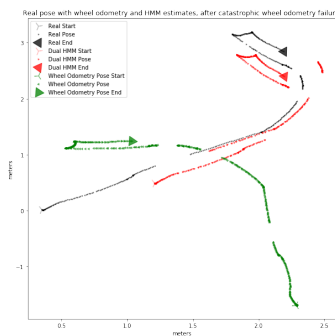


Fig. 24. Wheel Odometry and Dual-HMM estimates after Event

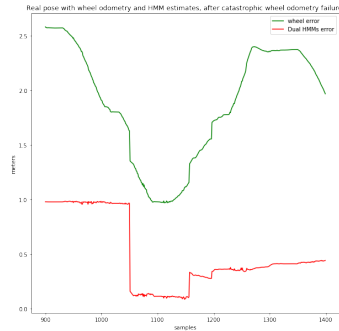


Fig. 25. Wheel Odometry and Dual-HMM error after Event

Also, although it was not shown here, the localization system remained operational during different home service tasks that required robot navigation in crowded environments. The experiments suggest that SLAM was performed successfully; it is important to notice that the online implementation of the algorithms made it possible for real-time optimization of the graph representation.

The same measurements were quantized combining different symbols from the same source using two clustering methods.

The same principle could also be applied to a wide variety of sensors or different features extracted from the same sensor measurements. This makes our system not only robust but also versatile, as it can be easily adapted to use a wide variety of sensors and action policies.

Future research is being conducted with aims in an active SLAM implementation using sensor fusion of laser, 2D image features, CNN extracted features, and 3D map features (octomaps [31]). At the same time, the optimal action policy to navigate the environment will be found using Q-learning or other appropriate reinforcement learning techniques.

References

1. Takashi Yamamoto, Tamaki Nishino, Hideki Kajima, Mitsunori Ohta, and Koichi Ikeda. Human {Support} {Robot} ({HSR}). In *{ACM} {SIGGRAPH} 2018 {Emerging} {Technologies}*, {SIGGRAPH} '18, pages 11:1—11:2, New York, NY, USA, 2018. ACM.
2. P. Zarchan and H. Musoff. *Fundamentals of Kalman Filtering: A Practical Approach*. Number v. 190 in *Fundamentals of Kalman filtering: a practical approach*. American Institute of Aeronautics and Astronautics, Incorporated.
3. R. Smith, M. Self, and P. Cheeseman. *Autonomous Robot Vehicles*. chapter Estimating, pages 167–193. Springer-Verlag, Berlin, Heidelberg, 1990.
4. Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard. Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. *IEEE Transactions on Robotics, Informatics and Automation*. 2022. Vol. 21 No. 1. ISSN 2713-3192 (print) 207 ISSN 2713-3206 (online) www.ia.spcras.ru

- 23(1):34–46, 2007.
5. D. Fox, W. Burgard, F. Dellaert, S. Thrun AAAI/IAAI, and undefined 1999.
 6. G. Grisetti, R. Kummerle, C. Stachniss, and W. Burgard. A Tutorial on Graph-Based {SLAM}. *Intelligent Transportation Systems Magazine, IEEE*, 2(4):31–43, 2010.
 7. M. Zaffar, S. Ehsan, R. Stolkin, and K.M. Maier. Sensors, slam and long-term autonomy: A review. In *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)*, pages 285–290, 2018.
 - Monte carlo localization: Efficient position estimation for mobile robots. *aaai.org*, 1999.
 8. Sebastian Thrun and Michael Montemerlo. The graph SLAM Algorithm with Applications to Large-Scale Mapping of Urban Structures. *The International Journal of Robotics Research*, 25(5-6):403–429, 2006.
 9. A. Soragna, M. Baldini, D. Joho, R. Kummerle, and G. Grisetti. Active slam using connectivity graphs as priors. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 340–346, 2019.
 10. João Machado Santos, David Portugal, and Rui P. Rocha. An evaluation of 2d slam techniques available in robot operating system. In *SSRR*, pages 1–6. IEEE, 2013.
 11. Giorgio Grisetti, Cyrill Stachniss, and Wolfram Burgard. Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. *IEEE Transactions on Robotics*, 23(1):34–46, 2007.
 12. Regis Vincent, Benson Limketkai, and Michael Eriksen. Comparison of indoor robot localization techniques in the absence of GPS. In Russell S Harmon, John H Holloway Jr., and J Thomas Broach, editors, *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XV*, volume 7664, pages 606–610. International Society for Optics and Photonics, SPIE, 2010.
 13. Kurt Konolige, Giorgio Grisetti, Rainer Kummerle, Wolfram Burgard, Benson Limketkai, and Regis Vincent. Efficient sparse pose adjustment for 2d mapping. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 22–29, 2010.
 14. G.O.S Ekhaguere. On notions of Markov property. *Journal of Mathematical Physics*, 18(11):2104–2107, 1977.
 15. Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. In *PROCEEDINGS OF THE IEEE*, pages 257–286, 1989.
 16. D. Fox, W. Burgard, F. Dellaert, S. Thrun AAAI/IAAI, and undefined 1999. Monte carlo localization: Efficient position estimation for mobile robots. *aaai.org*, 1999.
 17. Randall C. Smith and Peter Cheeseman. On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research*, 5(4), 1986.
 18. Savage J., Fuentes O. Map representation using hidden markov models for mobile robot localization. volume 161, 2018.
 19. A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, 39(1):1–38, 1977.
 20. Pierre Baldi and Yves Chauvin. Smooth {On}-{Line} {Learning} {Algorithms} for {Hidden} {Markov} {Models}. *Neural Comput.*, 6, 1993.
 21. Long Wen, Yang Zhao, Shuguang Li, Hong Cheng, and Chen Zhang. {MST}-{ResNet}: {A} {Multiscale} {Spatial} {Temporal} {ResNet} for {Steering} {Prediction}. pages 246–251, 2019.
 22. Jianbo Shi and Carlo Tomasi. Good Features to Track. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, 1994.
 23. Brendan J. Frey and Delbert Dueck. Clustering by passing messages between data points. *Science*, 315:972–977, 2007.

24. Ming Liang, Bin Yang, Yun Chen, Rui Hu, and Raquel Urtasun. Multi-task multi-sensor fusion for 3d object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
25. G.D. Forney. The viterbi algorithm. *Proceedings of the IEEE*, 61(3):268–278, 1973.
26. Agostino Martinelli. Modeling and {Estimating} the {Odometry} {Error} of a {Mobile} {Robot}. *IFAC Proceedings Volumes*, 34(6):407–412, 2001.
27. Yasuyoshi Yokokohji, Yoshihiro Kawai, Mizuho Shibata, Yasumichi Aiyama, Shinya Kotosaka, Wataru Uemura, Akio Noda, Hiroki Dobashi, Takeshi Sakaguchi, and Kazuhito Yokoi. World robot summit – summary of the pre-competition in 2018. *Advanced Robotics*, pages 1–24, 09 2019.
28. ROS. ROS (Robot Operating System), 2018.
29. Nathan Koenig and Andrew Howard. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2149–2154, Sendai, Japan, Sep 2004.
30. Imam Bukhori and Zool Ismail. Detection of kidnapped robot problem in monte carlo localization based on the natural displacement of the robot. *International Journal of Advanced Robotic Systems*, 14:172988141771746, 07 2017.
31. Kai Wurm, A Hornung, Maren Bennewitz, Cyrill Stachniss, and Wolfram Burgard. Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems. volume 2, 01 2010.
32. Bastian Steder, Christian Dornhege, Michael Ruhnke, Giorgio Grisetti, Cyrill Stachniss, and Alexander Kleiner. On Measuring the Accuracy of SLAM Algorithms. Technical report.

Fuentes Oscar — Team leader, Bio robotics laboratory, National Autonomous University of Mexico (UNAM). Research interests: machine learning, deep learning, computer vision, artificial intelligence, SLAM, navigation. The number of publications — 3. oscarfuentesfoto@gmail.com; 3000, Sirkuito Eskolar, Av. Universidad, 04510, Mexico City, Mexico; office phone: +52(55)56223041.

Savage Jesus — Ph.D., Professor, National Autonomous University of Mexico (UNAM). Research interests: autonomous mobile robots, digital signal processing, computer architectures. The number of publications — 0. robotssavage@gmail.com; 3000, Sirkuito Eskolar, Av. Universidad, 04510, Mexico City, Mexico; office phone: +52(55)56223041.

Contreras Luis — Ph.D., Research fellow, Advanced intelligence and robotics research center, Tamagawa University. Research interests: computer science, visual information, computer vision and robotics. The number of publications — 0. luis@lab.tamagawa.ac.jp; 6, Tamagawagakuen, 194-8610, Tokyo, Japan; office phone: +81 42-739-8111.

О. ФУЭНТЕС, Х. САВАЖ, Л. КОНТРЕРАС
**СИСТЕМА SLAM, ОСНОВАННАЯ НА СКРЫТЫХ МАРКОВСКИХ
МОДЕЛЯХ**

Фуэнтес О., Саваж Х., Контрерас Л. Система SLAM, основанная на скрытых марковских моделях.

Аннотация. Методы одновременной локализации и картографирования (SLAM) являются решением проблемы навигации сервисных роботов. Мы представляем графовую систему SLAM, основанную на скрытых марковских моделях (HMM), где показания датчиков представлены различными символами с использованием ряда методов кластеризации; затем символы объединяются в один для повышения точности с использованием двойных HMM. Универсальность нашей системы позволяет работать с датчиками разных типов или комбинировать датчики, а также реализовать активную или пассивную графовую систему SLAM. В подходе Graph-SLAM, предложенном Karto Robotics International в Cartographer, узлы представляют положение робота, а ребра представляют ограничения между ними. Узлы обычно задаются по непрерывным узлам, за исключением случаев обнаружения замыкания цикла, когда вводятся ограничения на несмежные узлы, что корректирует весь граф. Обнаружение цикливания не является тривиальным; в реализации ROS сопоставление сканирования выполняется с использованием регулировки положения разреженности (SPA). Картограф использует карту занятости, чтобы оценить положение, в котором карта отображается с помощью Gmapping. Робот Toyota HSR (Human Support Robot) использовался для создания набора данных как в реальных, так и в смоделированных условиях соревнований. В нашем представлении SLAM есть оценка одометрии колес в соответствии с начальным положением робота, 2D-лидарное сканирование Hokuyo для наблюдений, а также контроль сигналов и оценка карты окружающего пространства. Мы протестировали нашу систему в задаче о похищенном роботе, обучили начальную модель, затем улучшили ее в онлайн режиме и, наконец, решили задачу SLAM.

Ключевые слова: локализация, SLAM, навигация робота, картографирование, скрытая марковская модель, датчик.

Фуэнтес Оскар — руководитель группы, лаборатория биоробототехники, Национальный автономный университет Мексики (НАУМ). Область научных интересов: машинное обучение, глубокое обучение, компьютерное зрение, искусственный интеллект, SLAM, навигация. Число научных публикаций — 3. oscarfuentesfoto@gmail.com; Сиркуито Эсколар, проспект Универсидад, 3000, 04510, Мехико, Мексика; р.т.: +52(55)56223041.

Саваж Хесус — Ph.D., профессор, Национальный автономный университет Мексики (НАУМ). Область научных интересов: автономные мобильные роботы, цифровая обработка сигналов, компьютерная архитектура. Число научных публикаций — 0. robotssavage@gmail.com; Сиркуито Эсколар, проспект Универсидад, 3000, 04510, Мехико, Мексика; р.т.: +52(55)56223041.

Контрерас Луи — Ph.D., научный сотрудник, исследовательский центр передового интеллекта и робототехники, Университет Тамагава. Область научных интересов: информатика, визуальная информация, компьютерное зрение и робототехника. Число научных публикаций — 0. luis@lab.tamagawa.ac.jp; Тамагава Гакуэн, 6, 194-8610, Токио, Япония; р.т.: +81 42-739-8111.

Литература

1. Такаши Ямамото, Тамаки Нишино, Хидеки Кадзима, Мицунори Охта и Коичи Икеда.
2. П. Зарчан и Х. Мусофф. *Основы фильтрации Калмана: практический подход*. Number v. 190 в Основх калмановской фильтрации: практический подход. Американский институт аэронавтики и астронавтики, Incorporated.
3. Р. Смит, М. Селф и П. Чизмен. Автономные автомобили-роботы. глава Оценка, страницы 167–193. Шпрингер-Верлаг, Берлин, Гейдельберг, 1990.
4. Джорджио Гризетти, Сирилл Стахнисс и Вольфрам Бургард. Улучшенные методы построения сеток с помощью частиц Рао-Блэквелла Фильтры. *IEEE Transactions on Robotics*, 23 (1): 34–46, 2007.
5. D. Fox, W. Burgard, F. Dellaert, S. Thrun AAAI / IAAI и undefined 1999.
6. G. Grisetti, R. Kuemmerle, C. Stachniss и W. Burgard. Учебник по графическому SLAM . *Intelligent Transportation Systems Magazine, IEEE*, 2 (4): 31–43, 2010 г.
7. М. Заффар, С. Эхсан, Р. Столкин и К.М. Майер. Датчики, шум и длительная автономность: обзор. In *Конференция NASA / ESA 2018 по адаптивному оборудованию и системам (AHS)*, страницы 285–290, 2018. Локализация Монте-Карло: эффективное определение местоположения для мобильных устройств роботы. *aaai.org*, 1999.
8. Себастьян Трун и Майкл Монтемерло. Алгоритм graph SLAM с приложениями к крупномасштабному отображению Городские сооружения. *Международный журнал исследований робототехники*, 25 (5-6): 403–429, 2006.
9. А. Сорагна, М. Балдини, Д. Йохо, Р. Кюммерле и Г. Гризетти. Активный слэм с использованием графов связности в качестве априорных значений. In *Международная конференция IEEE / RSJ по интеллектуальным роботам, 2019 г. and Systems (IROS)*, страницы 340–346, 2019.
10. Жуан Мачадо Сантос, Д. Португалия и Руи П. Роча. Оценка методов двумерного удара, доступных в работе робота. В *SSRR*, страницы 1–6. IEEE, 2013.
11. Джорджио Гризетти, Сирилл Стахнисс и Вольфрам Бургард. Улучшенные методы построения сеток с помощью частиц Рао-Блэквелла Фильтры. *IEEE Transactions on Robotics*, 23 (1): 34–46, 2007.
12. Регис Винсент, Бенсон Лимкеткай и Майкл Эриксен. Сравнение методов локализации комнатных роботов при отсутствии GPS. In Russell S Harmon, John H Holloway Jr. и J Thomas Broach, редакторы, *Обнаружение и обнаружение мин, взрывоопасных предметов и скрытых объектов Targets XV*, том 7664, страницы 606–610. Международное общество оптики и фотоники, SPIE, 2010.
13. Курт Конолиге, Джорджио Гризетти, Райнер Кюммерле, Вольфрам Бургард, Бенсон Лимкеткай и Регис Винсент. Эффективная корректировка разреженной позы для 2d маппинга. На *Международной конференции IEEE / RSJ 2010 по интеллектуальным роботам and Systems*, страницы 22–29, 2010 г.
14. G.O.S. Ekhaguere. О понятиях марковского свойства. *Journal of Mathematical Physics*, 18 (11): 2104–2107, 1977.
15. Лоуренс Р. Рабинер. Учебное пособие по скрытым марковским моделям и избранным приложениям в распознавание речи. В *PROCEEDIES OF THE IEEE*, страницы 257–286, 1989.
16. D. Fox, W. Burgard, F. Dellaert, S. Thrun AAAI / IAAI и undefined 1999. Локализация Монте-Карло: эффективное определение местоположения для мобильных устройств роботы. *aaai.org*, 1999.

17. Рэндалл С. Смит и Питер Чизмен. О представлении и оценке пространственной неопределенности. *Международный журнал исследований робототехники*, 5 (4), 1986.
18. Сэвидж, Фуэнтес Представление карты с использованием скрытых марковских моделей для локализации мобильного робота. , том 161, 2018.
19. А.П. Демпстер, Н.М. Лэрд и Д.Б. Рубин. Максимальная вероятность получения неполных данных с помощью алгоритма EM. *ЖУРНАЛ КОРОЛЕВСКОГО СТАТИСТИЧЕСКОГО ОБЩЕСТВА, СЕРИЯ В*, 39 (1): 1–38, 1977.
20. Пьер Бальди и Ив Шовен. Smooth On-Line Learning Algorithms для Скрытых Марковских моделей . *Neural Comput.*, 6, 1993.
21. Лун Вэнь, Ян Чжао, Шугуан Ли, Хун Чэн и Чэнь Чжан. MST-ResNet : А Мульти-масштаб Spatial Temporal ResNet для рулевого управления
22. Джианбо Ши и Карло Томази. Хорошие возможности для отслеживания. em Конференция IEEE по компьютерному зрению и распознаванию образов, страницы 593-600, 1994.
23. Брендан Дж. Фрей и Делберт Дук. Кластеризация путем передачи сообщений между точками данных. em Science, 315: 972–977, 2007.
24. Мин Лян, Бинь Ян, Юн Чен, Жуй Ху и Ракель Уртасун. Многозадачная мультисенсорная комбинация для обнаружения трехмерных объектов. В em Proceedings of the IEEE / CVF Conference on Computer Vision and Распознавание образов (CVPR), июнь 2019 г.
25. Б-г Форни. Алгоритм Витерби. em Proceedings of the IEEE, 61 (3): 268–278, 1973.
26. Агостино Мартинелли. Моделирование и оценка ошибки одометрии мобильного робота }. em IFAC Proceedings Volumes, 34 (6): 407–412, 2001.
27. Ясуёси Ёкокодзи, Ёсихиро Каваи, Мидзухо Сибата, Ясумичи Айяма, Шинья Кото-сака, Ватару Уэмура, Акио Нода, Хироки Добаши, Такеши Сакагүти и Кадзухито Ёкои. World robot Summit - итоги предсоревнований 2018. em Advanced Robotics, страницы 1–24, 09, 2019.
28. РОС. ROS (операционная система роботов, 2018.
29. Натан Кенг и Эндрю Ховард. Разработка и использование парадигм для беседки, мульти-робота с открытым исходным кодом симулятор. In em Международная конференция IEEE / RSJ по интеллектуальным роботам и Systems, страницы 2149-2154, Сендай, Япония, сентябрь 2004 г.
30. Имам Бухори и Зоол Исмаил. Проблема обнаружения похищенного робота в локализации монте-карло на основе естественного перемещения робота. em Международный журнал передовых робототехнических систем, 14: 172988141771746, 07 2017.
31. Кай Вурм, А. Хорнунг, Марен Бенневиц, Сирилл Стахнисс и Вольфрам Бургард. Octomap: вероятностное, гибкое и компактное представление трехмерной карты для робототехнических систем. volume 2, 01 2010.
32. Бастиан Стедер, Кристиан Дорнхеге, Майкл Рунке, Джорджио Гризетти, Сирилл Стахнисс и Александр Клейнер. Об измерении точности алгоритмов SLAM. Технический отчет.

Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» <http://ia.spcras.ru/>. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTeX. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы разрешены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

Author guidelines

Interaction between each potential author and the Editorial board is realized through the pesoal account on the website of the journal "Informatics and Automation" <http://ia.spcras.ru/>. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTeX. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

Signed to print 01.02.2022

Printed in Publishing center GUAP, 190000, St. Petersburg, B. Morskaya 67, litera A, Russia

The journal is registered in the Russian Federal Agency for Communications and Mass-Media Supervision, certificate ПИ № ФС77-79228 dated September 25, 2020
Subscription Index П5513, Russian Post Catalog

Подписано к печати 01.02.2022. Формат 60×90 1/16. Усл. печ. л. 12,38. Заказ № 22.

Тираж 300 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП, 190000,
г. Санкт-Петербург, ул. Б. Морская, д. 67, лит. А

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций, свидетельство ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»