

ISSN 1816-0301

ИНФОРМАТИКА

2 (42)

АПРЕЛЬ – ИЮНЬ
2014

Редакционная коллегия:

Главный редактор

А.В. Тузиков

Заместитель главного редактора

М.Я. Ковалев

Члены редколлегии

С.В. Абламейко, В.В. Анищенко, П.Н. Бибило, М.Н. Бобов,
А.Н. Дудин, А.Д. Закревский, С.Я. Килин, В.В. Краснопрошин,
С.П. Кундас, Н.А. Лиходед, П.П. Матус, С.В. Медведев, А.А. Петровский,
Ю.Н. Сотсков, Ю.С. Харин, А.Ф. Чернявский, В.Н. Яролик
Н.А. Рудая (*заведующая редакцией*)

Адрес редакции:

220012, Минск,
ул. Сурганова, 6, к. 305
тел. (017) 284-26-22
e-mail: rio@newman.bas-net.by
<http://uiip.bas-net.by>

ИНФОРМАТИКА

ЕЖЕКВАРТАЛЬНЫЙ НАУЧНЫЙ ЖУРНАЛ

Издается с января 2004 г.

№ 2(42) • апрель-июнь 2014

СОДЕРЖАНИЕ

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

- Азаров И.С., Вашкевич М.И., Козлова С.В., Петровский А.А.** Система коррекции слуха на мобильной вычислительной платформе 5
- Гецэвіч Ю.С., Лабанаў Б.М., Пакладок Д.А.** Фанетычная і алафонная апрацоўка тэксту ў сінтэзатары беларускага і рускага маўлення для мабільных платформаў 25

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Шушкевич Г.Ч., Киселева Н.Н.** Экранирование звукового поля плоским упругим слоем и тонкой незамкнутой сферической оболочкой 36
- Швед О.Л.** Определение девиаторного сечения поверхности текучести при математическом моделировании упругопластического поведения материалов 49
- Переварюха А.Ю.** Разновидности аperiodической динамики в событийно-управляемой популяционной модели 58
- Бондоловский А.М.** Обзор моделей управления доходностью в гостиничном бизнесе 66

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

- Саечников А.В., Яцков Н.Н., Назаров П.В., Валлар Р., Апанасович В.В.** Анализ экспрессии генов в результате воздействия интерферона IFN- γ на клетку с использованием программного пакета GeneExpressionAnalyser 84
- Стрижнев А.Г., Ледник Г.В., Шихов А.А., Русакович А.Н.** Компенсация ошибки перехода цифровой следящей системы при круговом слежении 98

Хитриков С.В. Мобильная измерительная многоканальная система для проведения
стендовых и натурных испытаний..... 109

Воротницкий Ю.И., Кочин В.П., Стрикелев Д.А. Генетический алгоритм для
оптимизации структуры беспроводной сети с заданным качеством обслуживания 117

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

Ярмолик В.Н., Ярмолик С.В. Адресные последовательности для многократного
тестирования ОЗУ 124

Редактор Г.Б. Гончаренко
Корректор А.А. Михайлова
Компьютерная верстка Д.С. Гавинович

Сдано в набор 21.04.2014. Подписано в печать 26.05.2014.
Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс.
Усл. печ. л. 15,8. Уч.-изд. л. 15,5. Тираж 100 экз. Заказ 19.

Государственное научное учреждение «Объединенный институт проблем
информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1/274 от 04.04.2014.
ЛП № 02330/444 от 18.12.13.
Ул. Сурганова, 6, 220012, Минск.

© Объединенный институт проблем информатики
Национальной академии наук Беларуси, 2014

INFORMATICS

PUBLISHED QUATERLY

Issued since 2004

№ 2(42) • April-June 2014

CONTENTS

SIGNAL, IMAGE AND SPEECH PROCESSING

- Azarov E.S., Vashkevich M.I., Kozlova S.V., Petrovsky A.A.** Hearing correction system based on mobile computing platform..... 5
- Hetsevich Yu.S., Lobanov B.M., Pokladok D.A.** Phonetic and allophonic text processing in Belarusian and Russian speech synthesizer for mobile platforms..... 25

MATHEMATICAL MODELING

- Shushkevich G.Ch., Kiselyova N.N.** Sound field shielding by flat elastic layer and thin un-closed spherical shell..... 36
- Shwed O.L.** Defining deviatoric section of the fluidity surface for mathematical modeling of elastoplastic behavior of materials..... 49
- Perevaryukha A.Y.** Varieties of aperiodic dynamics in the event-driven population models 58
- Bandalouski A.M.** An overview of the hotel revenue management models 66

APPLIED INFORMATION TECHNOLOGIES

- Saetchnikov A.V., Yatskou M.M., Nazarov P.V., Vallar L., Apanasovich V.V.** Analysis of cellular reaction to IFN- γ stimulation by a software package GeneExpressionAnalyser 84
- Stryzhniou A.G., Lednik G.V., Shykhau A.A., Rusakovich A.N.** Compensation of the digital servosystem circular tracking transition error 98

Hitrikov S.V. Measuring mobile multi-channel system for bench and track tests	109
Vorotnitsky Y.I., Kochyn V.P., Strikelev D.A. A genetic algorithm for wi-fi network structure optimization under the given quality of service.....	117

LOGICAL DESIGN

Yarmolik V.N., Yarmolik S.V. Address sequences for multi run RAM testing	124
---	-----

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

УДК 681.84

И.С. Азаров¹, М.И. Вашкевич¹, С.В. Козлова², А.А. Петровский¹СИСТЕМА КОРРЕКЦИИ СЛУХА НА МОБИЛЬНОЙ
ВЫЧИСЛИТЕЛЬНОЙ ПЛАТФОРМЕ

Предлагается реализация цифрового слухового аппарата на основе мобильной вычислительной платформы iPhone. Рассматривается схема обработки сигнала с малой алгоритмической задержкой, позволяющая выполнять коррекцию слуха при нейросенсорной тугоухости и учитывающая особенности вычислительной платформы. Схема позволяет выполнять компенсацию потери слуха путем линейного частотно-зависимого усиления и широкополосной компрессии динамического диапазона сигнала. Для учета амплитудно-частотных характеристик динамиков гарнитуры аудиометрия выполняется непосредственно при помощи iPhone. Приводятся результаты экспериментального исследования реализованного слухового аппарата.

Введение

Тугоухость является одной из актуальных проблем в современном обществе и постоянно находится в центре внимания ведущих исследователей [1]. По данным Всемирной организации здравоохранения (ВОЗ) за 2013 г., около 360 млн человек в мире имеют умеренные или тяжелые нарушения слуха и их число к 2020 г. увеличится на 30 % [2].

Появление цифровых слуховых аппаратов (СА) позволило эффективно выполнять коррекцию такой сложной и распространенной патологии, как тугоухость. Методы и алгоритмы цифровой обработки сигналов, применяемые в современных СА, разнообразны и включают адаптивную фильтрацию, формирование диаграммы направленности, подавление эха и шумоподавление, компрессию динамического диапазона (нелинейное усиление), усиление частот в области пониженной чувствительности, подавление акустической обратной связи и др. [3].

Качественное расширение возможностей СА, произошедшее в последнее десятилетие, обусловлено развитием и наращиванием мощности портативных вычислительных устройств, увеличением их энергетического ресурса, а также более совершенными встроенными аналого-цифровыми и цифроаналоговыми преобразователями (АЦП и ЦАП соответственно). Конструкции СА становятся все более миниатюрными: на смену карманным аппаратам пришли встроенные в очки, затем заушные, внутриушные и внутриканальные. Традиционно СА проектируются в виде специальных устройств, единственной функцией которых является коррекция слуха. Узкая специализация обуславливает высокую стоимость данных устройств, а также создает сложности в обслуживании и настройке, требующие привлечения квалифицированных специалистов.

В последнее время широкое распространение (в том числе и среди людей с патологиями слуха) получили смартфоны – универсальные мультимедийные мобильные вычислительные платформы. Смартфоны имеют необходимые аппаратные средства для выполнения функций слухового аппарата: микрофон, АЦП, цифровой процессор, ЦАП, усилитель и динамики, причем возможно подключение внешних микрофона и динамиков в виде специальной гарнитуры. Работа программных приложений с аппаратными ресурсами смартфона обеспечивается при помощи операционной системы, для которой производителем предоставляется пакет инструментов разработчика. Устойчивое развитие данных платформ, их поддержка и малая стоимость обусловлены популярностью смартфонов и многочисленностью их пользователей. Идея, положенная в основу настоящей работы, заключается в том, чтобы исследовать возможность создания карманного СА на основе смартфона. В наиболее благоприятном случае, если аппаратные возможности смартфона соответствуют необходимым требованиям, создание СА ограничится разработкой специального программного обеспечения, которое может применять каждый пользователь с нарушениями слуха.

По сравнению с миниатюрными решениями карманный СА имеет следующие недостатки:
– более заметен;

– микрофон не располагается в ухе и поэтому не использует функциональные преимущества ушной раковины и естественной акустики наружного уха;

– не очень удобен для ношения.

Вместе с тем в силу своих конструктивных особенностей карманные СА имеют ряд важных преимуществ:

– большое расстояние между микрофоном и динамиком препятствует возникновению акустической обратной связи (АОС), что позволяет использовать большое акустическое усиление и более простой алгоритм обработки (в миниатюрных СА обязательна реализация схемы подавления АОС);

– большие физические размеры позволяют реализовать удобные функции управления для людей с плохой моторикой;

– возможно использование различных наушников в зависимости от патологии и предпочтений пациента (в том числе больших в виде накладных чашек, не приводящих к возникновению эффекта окклюзии);

– благодаря большим динамикам и большому ресурсу батарей можно достигнуть наибольшего уровня звукового давления и получить высокое качество звука;

– в отличие от внутриканальных СА карманные не выходят из строя из-за попадания в них серы или влаги;

– при использовании динамика с костной проводимостью отсутствует механическая обратная связь. Такой динамик плотно примыкает к височной кости и, создавая бесшумные вибрации, способен передавать их в среднее ухо, минуя барабанную перепонку.

Благодаря тому что СА реализованы на популярной мобильной платформе, а не в виде отдельного узкоспециализированного устройства, для них характерны и некоторые дополнительные преимущества:

– большой доступный ресурс энергии и возможность использовать сложные алгоритмы обработки и высокую частоту дискретизации;

– широкое распространение, многофункциональность и высокий спрос обуславливают устойчивую поддержку и развитие этих устройств со сторон производителей;

– программная гибкость, т. е. возможность программной реализации произвольного алгоритма обработки сигналов;

– доступность средств разработчика и развитая система распространения программных продуктов для данных устройств;

– возможность применения алгоритмов коррекции слуха не только к акустическим сигналам, но и к аудио- и видеозаписям, телефонным звонкам и т. д.;

– индивидуальная настройка аппарата для конкретного пациента в простых случаях не требует специального оборудования и квалификации и может выполняться самим пользователем;

– возможность удаленной настройки аппарата специалистом на основе диагноза пациента;

– наличие встроенной видеокамеры дает теоретическую возможность обработки комбинированной (аудио- и видео-) информации для повышения разборчивости речи;

– возможность использования различных гарнитур с разными характеристиками микрофонов и наушников. Возможна разработка специальной гарнитуры с расположением микрофона в ушной раковине для использования звуковых характеристик наружного уха;

– заметность слухового аппарата не вызывает психологических неудобств, поскольку смартфон не ассоциируется у окружающих с какими бы то ни было слуховыми патологиями;

– для пользователей смартфонов с нарушением слуха нет необходимости приобретать и носить отдельное устройство.

1. Принципы построения современных цифровых слуховых аппаратов

1.1. Субполосная обработка сигнала

Основной функцией СА является частотно-зависимое усиление входных сигналов для компенсации нарушений слуха [3]. Процесс обработки сигнала может включать вспомогательные функции, такие как подавление АОС, подавление шума и компрессия динамического диапазона. Все перечисленные функции, за исключением подавления АОС, могут быть выполнены

путем субполосной декомпозиции и отдельного усиления каждого субполосного сигнала. На рис. 1 показана соответствующая схема обработки [4]. Входной звуковой сигнал $v(n)$ регистрируется микрофоном и раскладывается при помощи банка фильтров анализа (БФА) на субполосные сигналы с пониженной частотой дискретизации $u(m)$. Временные индексы исходного и субканальных сигналов обозначены n и m соответственно (контур, работающий на пониженной частоте, выделен жирными линиями на схеме).

Отдельное усиление каждого субполосного сигнала выполняется умножением на коэффициенты $g(m)$, которые вычисляются для каждого момента времени исходя из уровня входного и выходного сигнала, энергии шума и частотного профиля потери слуха. При этом блоком усиления выполняется сразу несколько функций: частотная коррекция потери чувствительности слуха; компрессия (компенсация функции нелинейного усиления улитки); шумоподавление на основе спектрального взвешивания. Обработанный широкополосный сигнал синтезируется при помощи банка фильтров синтеза (БФС). Широкополосный сигнал умножается на общий коэффициент усиления g_c , который обеспечивает комфортный уровень прослушивания (параметр g_c обычно регулируется пользователем при помощи внешнего регулятора).

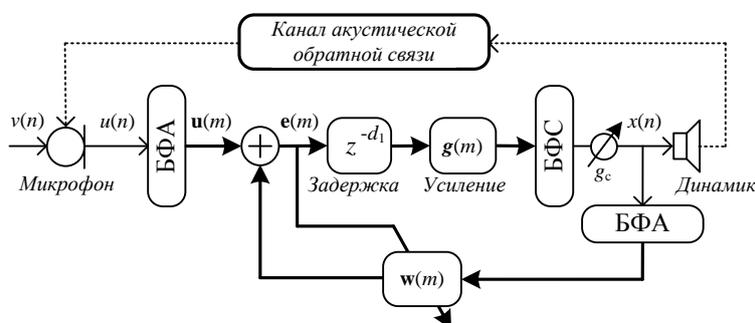


Рис. 1. Общая схема обработки сигнала в современном цифровом слуховом аппарате

В миниатюрных СА максимальное доступное усиление сигнала ограничено появлением АОС между микрофоном и динамиком. Для подавления АОС чаще всего используется схема адаптивной фильтрации. Коэффициенты адаптивного фильтра $w(m)$ постоянно обновляются таким образом, чтобы минимизировать энергию сигнала-остатка $e(m)$. Обновление коэффициентов выполняется во временной либо частотной области с помощью какой-либо модификации метода наименьших квадратов (LMS) [5, 6]. Задержка z^{-d_1} смещает сигналы друг относительно друга и влияет на оценку коэффициентов адаптивного фильтра. Большие значения d_1 обеспечивают более сильное подавление АОС и повышение возможной степени усиления сигнала [4]. Преимуществом схемы с субканальной декомпозицией сигнала является относительно быстрая сходимость адаптивного фильтра и сокращение вычислительных ресурсов за счет понижения частоты дискретизации [7].

Компрессия сигнала выполняется с целью компенсации утраченной функции улитки внутреннего уха выполнять нелинейное усиление. Компрессия, применяемая для коррекции, должна иметь очень короткое время активации [2]. При использовании таких малых временных интервалов происходит быстрое изменение уровня сигнала, что приводит к появлению низкочастотных искажений. Дополнительным преимуществом приведенной выше схемы является возможность снизить уровень этих искажений, используя в низкочастотной части сигнала более продолжительное время активации.

1.2. Сокращение задержки сигнала в слуховом аппарате

Одной из наиболее важных характеристик СА является задержка обработки сигнала. Временная задержка вызывает окрашивание звука собственного голоса пользователя СА. Во время разговора голос быстро достигает улитки посредством костной проводимости и

складывается с задержанным и усиленным звуком из СА. Этот эффект становится более заметным при использовании больших отверстий в ушном вкладыше. В работах [8, 9] показано, что даже такие малые задержки обработки сигнала, как 4–8 мс, ощущаются пациентом и снижают субъективное качество звука. В СА невозможно избежать задержек, обусловленных наличием АЦП и ЦАП, которые составляют от 0,4 до 2 мс в зависимости от реализации [10]. В условиях создания СА на базе смартфона длительность неустранимой аппаратной задержки будет существенно больше из-за внутренней аппаратной и программной реализации сигнальных конвейеров, поэтому в данном случае требуется схема обработки сигнала с минимально возможной алгоритмической задержкой. Общая алгоритмическая задержка классической схемы с субканальной декомпозицией, приведенной выше, является нежелательно большой [11–13] из-за последовательных элементов, включенных в путь прямого распространения сигнала: БФА, БФС и задержки z^{-d_1} , которая используется для стабилизации системы подавления АОС. В работах [4, 14] предлагаются схемы с субканальным усилением сигнала без БФС. Для этого используются полосовые фильтры (параметрические фильтры частотной компенсации [4] либо параметрические кохлеарные фильтры [14]), выходы которых суммируются (рис. 2).

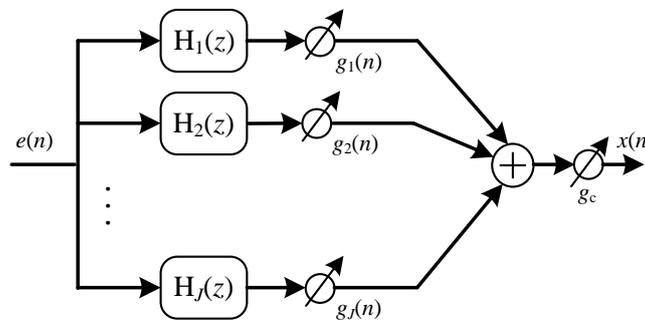


Рис. 2. Метод субканальной коррекции амплитудного спектра с малой задержкой

В работах [14, 15] используется субканальная система подавления АОС, не вносящая дополнительных задержек в прямой канал. Перечисленные решения позволяют существенно сократить общую алгоритмическую задержку, однако отсутствие БФС делает невозможным понижение частоты дискретизации субполосных сигналов и требует больших вычислительных затрат. Кроме того, полосовые фильтры с бесконечными импульсными характеристиками, используемые в работах [4, 14], имеют нелинейную фазочастотную характеристику (ФЧХ) и для их согласования используется оптимизационная процедура, которая в зависимости от целевого амплитудного профиля подбирает субканальные коэффициенты усиления. Данная процедура также является достаточно затратной в вычислительном плане. Еще одной важной особенностью данной схемы является то, что обработка сигнала ведется отсчет за отсчетом. Это предполагает использование специального аппаратного вычислителя.

1.3. Обработка сигнала в слуховом аппарате на базе мобильной платформы

Существующие мобильные платформы позволяют обрабатывать сигнал в реальном времени только отдельными фреймами длиной не менее 6 мс. В этих условиях желательно, чтобы алгоритм обработки позволял применять блочные операции. Учитывая перечисленные особенности, предлагается использовать широкополосную (без субполосной декомпозиции) схему коррекции слуха (рис. 3).



Рис. 3. Предлагаемая схема обработки сигнала

Схема использует широкополосный способ коррекции огибающей спектра при помощи фильтра с конечной импульсной характеристикой (КИХ). Частотная характеристика фильтра формируется на основе аудиограммы пользователя таким образом, чтобы выполнять усиление частот, к которым наблюдается ослабление чувствительности. Для управления уровнем громкости используются регулятор чувствительности микрофона g_m и регулятор выходного уровня g_c . Оба регулятора являются внешними и доступны пользователю для оперативного использования в зависимости от акустической обстановки. Коэффициенты усиления $g(n)$ изменяются таким образом, чтобы обеспечивать компрессию сигнала в соответствии с заданным уровнем потери слуха.

Учитывая, что смартфон использует стереогарнитуру с двумя динамиками, можно изменить схему с бинауральной коррекцией слуха, выполняющей обработку сигнала отдельно для левого и правого уха (рис. 4).

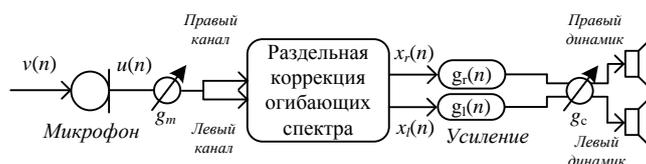


Рис. 4. Предлагаемая бинауральная схема обработки сигнала с одним микрофоном

Прямая реализация предлагаемой широкополосной схемы исключает использование шумоподавления на основе спектрального взвешивания, поскольку в ней не выполняется частотно-временное преобразование. Схема не содержит блока подавления АОС, поскольку при ее реализации на смартфоне предполагается использовать микрофон и динамик на большом расстоянии друг от друга. Благодаря тому что предлагаемая широкополосная схема не содержит БФА и БФС, а также стабилизирующего элемента задержки для подавления АОС, в ней достигается минимально возможная алгоритмическая задержка, равная групповой задержке фильтра коррекции спектральной огибающей.

2. Определение целевых параметров звукоусиления

Различают несколько составляющих тугоухости: ухудшение звуковосприятия, нарушения ощущения громкости и бинауральное снижение слуха [3]. Для оценки характеристик каждого из этих компонентов используют специальные методики и средства. Характер и степень ухудшения восприятия определяются при помощи аудиометрии, нарушения ощущения громкости оцениваются при помощи методов категоризации громкости, а бинауральные нарушения – путем тестирования различий разборчивости в каждом из ушей. В СА, реализованном на мобильной платформе с использованием предлагаемой схемы, возможно выполнить компенсацию ухудшения звуковосприятия при помощи фильтра коррекции огибающей, а компенсацию нарушения ощущения громкости (рекруитмента) – при помощи элемента динамического усиления.

2.1. Аудиограмма

При подборе и настройке СА аудиометрия дает самую первую и основную характеристику нарушения слуха. В результате аудиометрии формируется таблица минимальных уровней восприятия (англ. hearing threshold level, НТЛ) тональных сигналов, называемая аудиограммой. Пороги восприятия определяются на частотах 125, 250, 500, 1000, 2000, 4000 и 8000 Гц и могут принимать значения от -10 до 120 дБ. В 1997 г. ВОЗ была утверждена единая Международная классификация степеней тугоухости, которая чаще всего используется на практике. Для определения степени тугоухости учитываются усредненные значения порогов звуковосприятия на четырех частотах: 500, 1000, 2000 и 4000 Гц. Исходя из этой классификации, нормальным считается слух при порогах слышимости на речевых частотах, не превышающих 10 дБ. Легкое нарушение слуха регистрируют при порогах слышимости $10-25$ дБ. Потеря слуха свыше 90 дБ определяется как глухота.

Асимметричное нарушение слуха классифицируется по лучше слышащему уху.

Выделяют четыре степени тугоухости:

I степень – снижение слуха в пределах 26–40 дБ (человек с такой потерей слуха с трудом распознает негромкую речь, но чувствует себя уверенно в тихой обстановке);

II степень – 41–55 дБ (имеются трудности в понимании беседы, особенно когда присутствует шум на заднем плане. Повышенная громкость необходима для ТВ и радио);

III степень – 56–70 дБ (значительно задета чистота речи. Речь должна быть громкой, возможны трудности при групповой беседе);

IV степень – 71–90 дБ (имеется значительная потеря слуха – человек не слышит нормальную разговорную речь, испытывает трудности при распознавании даже громкой речи, способен понимать крик и преувеличенно четкую и громкую речь).

Аудиограмма дает возможность выполнить расчет необходимого (целевого) усиления и максимальной выходной мощности, обеспечивающих для пользователя максимально комфортный уровень громкости речевого сигнала при настройке обычного СА. Снятие аудиограммы пользователя выполняется в лабораторных условиях в помещении со звукоизоляцией и использованием специального оборудования, обеспечивающего генерацию и усиление тональных сигналов заданной частоты. Снятием аудиограммы руководит квалифицированный медицинский работник, создавая постепенно нарастающее звуковое давление в каждом ухе пациента. Как только звук становится различимым, пациент реагирует нажатием на сигнальную кнопку.

При помощи мультимедийной платформы можно автоматизировать весь процесс таким образом, чтобы пользователь мог самостоятельно выполнить аудиометрию. Для этого используются программные генераторы тональных сигналов и интерфейсная кнопка для реакции на превышение порога восприятия. Тональный сигнал описывается функцией косинус с нарастающей амплитудой и постоянной частотой:

$$s(n) = A(n) \cos\left(2\pi \frac{f}{f_s} n\right),$$

где $A(n)$ – изменяющаяся (нарастающая) амплитуда; f – заданная частота, Гц; f_s – частота дискретизации. Скорость нарастания амплитуды определяется продолжительностью сигнала и доступным динамическим диапазоном ЦАП устройства. Нормированную амплитуду (единица соответствует максимально возможному значению) можно рассчитать по следующей формуле:

$$A(n) = 10^{(b-1)\left(\frac{n}{f_s l} - 1\right) \log_{10} 2},$$

где b – разрядность ЦАП; l – длительность сигнала, с. При таком способе расчета уровень сигнала в децибелах растет линейно от минимально возможного значения до максимального:

$$\text{level}(n) = 20 \log_{10}(A(n)) = 20(b-1) \left(\frac{n}{f_s l} - 1 \right) \log_{10} 2.$$

Для снятия аудиограммы в лабораторных условиях используются специальные наушники с ровной амплитудно-частотной характеристикой (АЧХ). Для определения требуемого усиления в СА на каждой контрольной частоте учитывается частотная характеристика динамика СА, ушного канала и размер отверстия в ушном вкладыше. В случае выполнения измерений при помощи смартфона условия аудиометрии и эксплуатации являются одинаковыми и упрощенно можно считать, что неровности в АЧХ динамиков гарнитуры являются потерей слуха, и выполнить расчет целевых коэффициентов усиления без вносимых обычно поправок.

2.2. Расчет целевого усиления и коэффициентов корректирующего фильтра

Одним из наиболее известных способов расчета целевого усиления является правило полуусиления, согласно которому линейное усиление должно примерно равняться половине уровня потери слуха на каждой частоте для того, чтобы обеспечивать комфорт при прослушивании и речевую разборчивость [3]. Тем не менее длительное время основным методом, использовавшимся в клинической практике, был сравнительный перебор, при котором пациент пробовал разные СА и выбирал тот, который обеспечивал наилучшие результаты. В исследовательской работе, опубликованной в 1975 г., был сделан вывод, что АЧХ СА, определяющая уровень усиления в зависимости от частоты, существенно влияет на повышение разборчивости речи для людей с сенсоневральной тугоухостью. Это исследование инициировало постепенный процесс перехода от метода подбора к использованию директивных расчетных формул. Прежде чем появились нелинейные слуховые аппараты, методы расчетов частотно-зависимого линейного усиления представляли собой вариации вышеприведенного правила полуусиления.

Сегодня существует несколько классических методов расчета целевого усиления на основе аудиограммы: Berger [16], POGO (Prescription of gain and output – предписание усиления и выхода) [17] и NAL-R (National Acoustic Laboratories, Australia – Национальные акустические лаборатории, Австралия) [18]. Соответствующие расчетные формулы приведены в табл. 1 [3], HTL X обозначает измеренные значения порогов слышимости на частоте X .

Таблица 1

Расчетные формулы целевого усиления

Частота, Гц	Berger	POGO	NAL-R $X = 0,05(\text{HTL}500 +$ $+ \text{HTL}1k + \text{HTL}2k)$
250		$\text{HTL}250/2 - 10$	$X + 0,31\text{HTL}250 - 17$
500	$\text{HTL}500/2$	$\text{HTL}500/2 - 5$	$X + 0,31\text{HTL}500 - 8$
750	$\text{HTL}750/1,8$	$\text{HTL}750/2 - 3$	$X + 0,31\text{HTL}750 - 3$
1000	$\text{HTL}1k/1,6$	$\text{HTL}1k/2$	$X + 0,31\text{HTL}1k - 1$
1500	$\text{HTL}1,5k/1,6$	$\text{HTL}1,5k/2$	$X + 0,31\text{HTL}1,5k - 1$
2000	$\text{HTL}2k/1,5$	$\text{HTL}2k/2$	$X + 0,31\text{HTL}2k - 1$
3000	$\text{HTL}3k/1,7$	$\text{HTL}3k/2$	$X + 0,31\text{HTL}3k - 2$
4000	$\text{HTL}4k/1,9$	$\text{HTL}4k/2$	$X + 0,31\text{HTL}4k - 2$
6000	$\text{HTL}6k/2$	$\text{HTL}6k/2$	$X + 0,31\text{HTL}6k - 2$

Все три перечисленных метода расчета могут быть использованы для расчета целевого усиления в программной реализации СА. Выбор наиболее подходящего из них делается либо самим пользователем исходя из его предпочтений, либо специалистом на основе характера патологии и лабораторных измерений.

Для тяжелой и глубокой степени потери слуха метод NAL-R был переработан. Модифицированный вариант метода, называемый NAL-RP [19], позволяет получить большее усиление и более равномерную частотную характеристику, чем NAL-R. Коэффициент X вычисляется по формуле

$$X = \begin{cases} 0,05(\text{HTL}500 + \text{HTL}1k + \text{HTL}2k), & \text{если } (\text{HTL}500 + \text{HTL}1k + \text{HTL}2k) < 180 \text{ дБ}; \\ 0,05(\text{HTL}500 + \text{HTL}1k + \text{HTL}2k), & \text{если } (\text{HTL}500 + \text{HTL}1k + \text{HTL}2k) \geq 180 \text{ дБ}. \end{cases}$$

Если порог слышимости на частоте 2 кГц превышает 95 дБ, к полученным значениям усиления нужно прибавить поправочные коэффициенты (табл. 2).

Таблица 2

Поправочные коэффициенты при тяжелых формах тугоухости

Пороги слышимости на частоте 2 кГц	Частота, Гц								
	250	500	750	1000	1500	2000	3000	4000	6000
95	4	3	1	0	-1	-2	-2	-2	-2
100	6	4	2	0	-2	-3	-3	-3	-3
105	8	5	2	0	-3	-5	-5	-5	-5
110	11	7	3	0	-3	-6	-6	-6	-6
115	13	8	4	0	-4	-8	-8	-8	-8
120	15	-	4	0	-5	-9	-9	-9	-9

В соответствии с полученными значениями целевого усиления рассчитываются коэффициенты фильтра коррекции спектральной огибающей. В данной реализации СА используются фильтры с линейной ФЧХ, что позволяет достигать равной временной задержки на всех частотах обрабатываемого сигнала вне зависимости от целевых значений усиления. Данное свойство особенно важно в бинауральном режиме коррекции слуха, поскольку различия в групповых задержках фильтров для правого и левого каналов вызывают фазовые смещения и как следствие, искажения пространственного восприятия звука. Фильтры с линейной ФЧХ удобно синтезировать при помощи оконного метода [20]. Алгоритм получения коэффициентов фильтра кратко приведен ниже.

1. Целевая частотная характеристика фильтра формируется путем интерполяции целевых значений усиления в точках, соответствующих частотным отсчетам. Фаза всех частотных компонентов принимается нулевой.

2. Вычисляется обратное преобразование Фурье целевой частотной характеристики.

3. Выполняется циклический сдвиг полученного вектора значений на половину числа его элементов.

4. Окончательная импульсная характеристика фильтра находится путем умножения полученных значений на оконную функцию.

Число коэффициентов фильтра выбирается исходя из частоты дискретизации сигнала и допустимой групповой задержки.

2.3. Компенсация нарушения ощущения громкости

Установлено, что тихие звуки усиливаются здоровой улиткой на 50–60 дБ, а громкие не усиливаются. Изменение коэффициента усиления физиологически обеспечивается наружными волосковыми клетками. При потере слуха более чем на 60 дБ волосковые клетки теряют свои регулятивные возможности и характеристика улиткового усилителя превращается в линейную (рис. 5) [3]. В этом случае требуется компрессия с коэффициентом 2,5:1. Слышимый динамический диапазон пациента сужается, и пациент может относительно хорошо воспринимать громкие звуки по сравнению с тихими. Другими словами, при увеличении уровня сигнала восприятие громкости при тугоухости приближается к нормальному. Это требует некоторого увеличения громкости, которое называется рекруитментом.

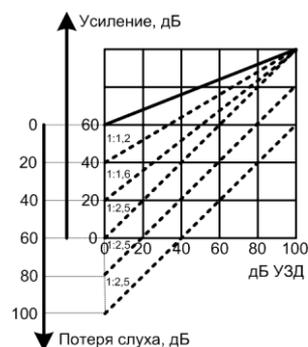


Рис. 5. Усиление в улитке

Основным назначением компрессора является усиление тихих звуков и в то же время – предохранение от чрезмерного усиления громких звуков. Общая схема (рис. 6) показывает функциональную организацию простого компрессора. Сначала измеряется текущий уровень сигнала, а затем вычисляются требуемые коэффициенты усиления для данного уровня $g(n)$ исходя из степени потери слуха.

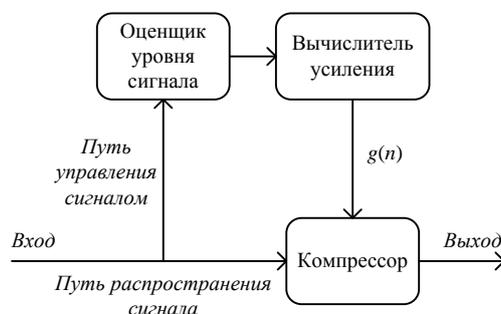


Рис. 6. Простая схема компрессии сигнала

Компрессор, схема которого изображена на рис. 6, называется компрессором с прямой связью, поскольку в нем выполняется оценка уровня входного сигнала перед выполнением компрессии. Если же выполняется оценка уровня выходного сигнала после компрессии, то такая схема называется схемой с обратной связью. В СА преимущественно используются компрессоры с прямой связью.

2.4. Временные параметры компрессора динамического диапазона

Динамические характеристики компрессора определяются скоростью реакции оценщика уровня сигнала на изменения уровня сигнала. Для того чтобы избежать появления артефактов, изменения коэффициентов усиления не должны быть резкими. Это достигается сглаживанием действительной амплитудной огибающей сигнала. Системы компрессии с сильным сглаживанием могут использоваться либо при частом изменении акустической обстановки, либо с целью выравнивания общего уровня прослушивания. Системы с коротким временем реакции имеют своей целью регулировку усиления перед каждой новой фоновой или слоговой обрабатываемой речью, в этом случае компрессор часто называют фоновым или слоговым. Ни один оценщик уровня не является идеальным, всегда существует некоторая задержка между реальным уровнем сигнала и его оценкой. Быстрота реакции оценщика уровня определяется временными константами.

Оценка уровня сигнала $x(n)$ в заданный момент времени t_n обозначается Φ_{t_n} и обновляется с заданным фиксированным временным интервалом $\Delta t = t_n - t_{n-1}$ [21]. Обычно Φ_{t_n} вычисляется как взвешенное среднее между его предыдущим значением $\Phi_{t_{n-1}}$ и средним квадратов отсчетов сигнала $\bar{x}^2_{[t_{n-1}, t_n]}$, находящихся в интервале между t_{n-1} и t_n :

$$\Phi_{t_n} = a\Phi_{t_{n-1}} + (1-a)\bar{x}^2_{[t_{n-1}, t_n]},$$

где $a \in [0, 1]$. Выбор константы a определяет скорость реакции оценщика уровня. Если a близко к 1, оценка уровня меняется очень медленно, в то время как при приближении τ к 0 время реакции уменьшается.

Еще одна временная константа, обозначенная как τ , используется для настройки параметров компрессора и принимает различные значения в зависимости от того, повышается или понижается уровень сигнала. Если уровень сигнала повышается, то она называется временем срабатывания, а если понижается, то временем восстановления. В соответствии со стандартом ИЕС 60118-2 время срабатывания, обозначаемое как α , – это время, которое необходимо для то-

го, чтобы выход достиг окончательного значения (с точностью ± 2 дБ) при резком увеличении уровня звукового давления (УЗД) сигнала от 55 до 80 дБ. Аналогично время восстановления, обозначаемое как β , является временем, затрачиваемым на установление окончательного значения (с ± 2 дБ) при резком изменении уровня сигнала от 80 до 55 дБ УЗД. В обоих случаях предполагается, что выходной уровень достиг устойчивого значения раньше резкого изменения уровня.

Допустим, что сигнал изменяется от x_1 до x_2 в момент времени t_1 , т. е. $x(t_n) = x_1$ для $n \leq 0$ и $x(t_n) = x_2$ для $n > 0$ и $\Phi_{t_0} = x_1^2$. Тогда

$$\Phi_{t_n} = a^n (x_1^2 - x_2^2) + x_2^2.$$

Допустим, что требуемый выходной уровень достигается в момент времени t_{n_t} , где τ означает временную константу (время срабатывания либо восстановления). Тогда τ можно вычислить из заданного значения веса a :

$$\tau = n_t \Delta t = \frac{\ln(x_2^2 - \Phi_{t_{n_t}}) - \ln(x_2^2 - x_1^2)}{\ln a} \Delta t.$$

Искомый весовой коэффициент вычисляется из временной константы τ следующим образом:

$$a = \left(\frac{x_2^2 - \Phi_{t_{n_t}}}{x_2^2 - x_1^2} \right)^{\Delta t / \tau}.$$

При вычислении весового коэффициента для заданного времени срабатывания значения x_1^2 и x_2^2 устанавливаются равными значениям, соответствующим 55 и 80 дБ УЗД; при вычислении весового коэффициента для заданного времени восстановления значения x_1^2 и x_2^2 устанавливаются равными значениям, соответствующим 80 и 55 дБ УЗД. Преимуществом подхода использования временной константы является то, что он остается универсальным вне зависимости от частоты дискретизации сигнала и от скорости изменения уровня сигнала.

Выбор временной константы в компрессоре СА является спорным вопросом. Основная цель быстрой компрессии – это регулировка усиления перед каждым звуком речи для того, чтобы тихие фонемы (согласные) получали более высокий уровень усиления, чем громкие. В опубликованных научных результатах не было показано, что компрессия сигнала повышает разборчивость по сравнению с линейным способом усиления, обеспечивающим постоянно комфортный для прослушивания уровень сигнала. Однако в том случае когда речь имеет широкий динамический диапазон, компрессия обладает заметным преимуществом. Некоторые исследования показали, что пациенты со слуховыми патологиями предпочитают качество звука, обеспечиваемое системой компрессии с медленной реакцией, особенно в шумной обстановке [21]. Одна из проблем медленной компрессии – недостаточное усиление тихих звуков, следующих сразу за громкими, поскольку компрессору нужно продолжительное время для регулировки усиления. Это может вызывать у пользователя СА пропуски фрагментов речи после громких звуков. Напротив, проблемой при использовании быстрой компрессии является то, что тихий шум становится слышимым на тихих фрагментах либо паузах речи. Показано, что при наличии статического фонового шума быстрая компрессия предпочтительнее медленной. Согласно данным, приведенным в работе [3], время активации должно быть коротким для того, чтобы обеспечивать слоговую компрессию (время срабатывания меньше 10 мс и время восстановления меньше 50 мс).

3. Реализация слухового аппарата на мобильной платформе

3.1. Схема включения приложения

В качестве мобильной платформы используется смартфон компании Apple семейства iPhone 5. Для реализации СА необходимо выполнять обработку в реальном масштабе времени с малой алгоритмической задержкой. Программное приложение должно установить потоковую передачу цифрового сигнала от микрофона к динамику. Внутренняя организация функциональных блоков iPhone для обработки звуковых сигналов позволяет применить алгоритм коррекции слуха к сигналу, формируемому на выходе микшера (рис. 7). Микшер отвечает за синхронизацию и захват сигналов из микрофонной линии и программных приложений системы. В результате можно выполнять коррекцию не только микрофонного сигнала, но и сигналов различных программных приложений смартфона, например аудио- и видеопроигрывателей.

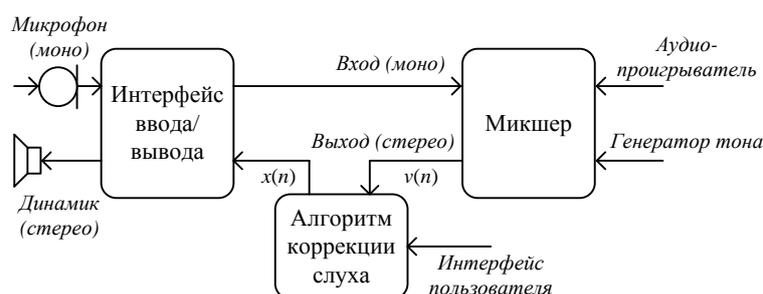


Рис. 7. Схема включения приложения коррекции слуха в звуковую систему iPhone

Приложение для коррекции слуха имеет два режима: аудиометрии и коррекции. В режиме аудиометрии измеряются пороги слышимости путем формирования тональных сигналов с нарастающей амплитудой. В режиме коррекции выполняется обработка сигнала с учетом полученных порогов. Пользователь может подобрать параметры для достижения наилучшей разборчивости.

3.2. Полоса пропускания

Полоса пропускания в современных СА не превышает 10 кГц, поскольку используемая частота дискретизации составляет 16–20 кГц и существенная часть слышимого звукового спектра не обрабатывается. Ограничение частоты вызвано тем, что используемые в СА специализированные цифровые процессоры обрабатывают сигнал с частотой, в несколько раз превышающей частоту дискретизации, что ведет к увеличению потребляемой мощности и сокращению срока службы батарейки. Поэтому перед разработчиками СА всегда стоит выбор: расширить полосу пропускания либо увеличить срок службы элементов питания. Учитывая, что платформа iPhone имеет перезаряжаемый элемент питания и ориентирована на мультимедийные приложения, возможно использование частоты дискретизации 44,1 кГц, обеспечивающей наилучшее качество звука.

3.3. Алгоритм обработки сигнала в слуховом аппарате на мобильной платформе

Вычислительная платформа iPhone позволяет обрабатывать входной сигнал отдельными кадрами по R отсчетов. Размер кадра, как правило, колеблется в пределах от 128 до 1024 отсчетов и выбирается равным степени числа 2. Для реализации функции СА с целью уменьшения задержки лучше выбирать размер кадра равным 128 или 256. Обработка сигнала заключается в выполнении двух основных процедур: коррекции огибающей спектра (на основе аудиограммы пользователя) и динамической компрессии сигнала с целью компенсации нарушения ощущения громкости.

Коррекция огибающей спектра может быть выполнена при помощи КИХ-фильтра с фиксированной частотной характеристикой. Фильтр синтезируется заранее по индивидуальной аудиограмме пользователя с использованием правил расчета целевого усиления POGO, NAL-R или Berger [3]. Для уменьшения вычислительной сложности фильтрацию необходимо выпол-

нять в частотной области. Для этой цели предлагается использовать метод перекрытия с суммированием [20].

На вход алгоритма обработки сигнала (рис. 8) поступает дискретизированный сигнал $x[n]$, который разбивается на кадры $x^{(m)}[r]$:

$$x^{(m)}[r] = x[mR + r],$$

где m – номер кадра; $r = 1, \dots, R$ – индекс отсчета внутри кадра.

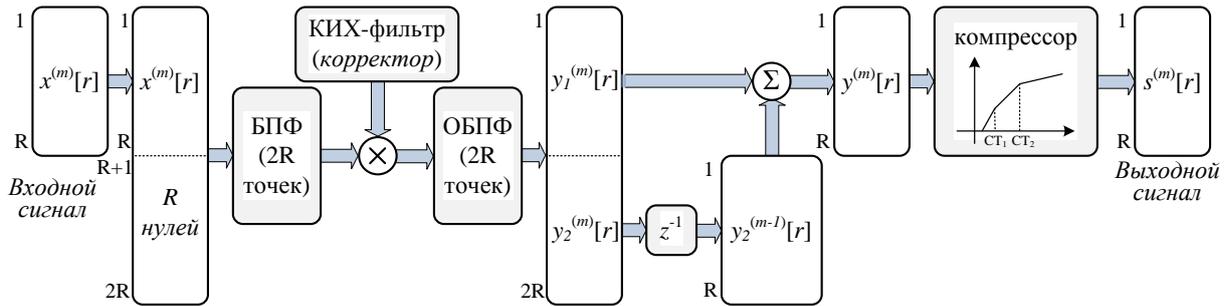


Рис. 8. Предлагаемый алгоритм обработки сигнала в слуховом аппарате на базе мобильной платформы

Для выполнения линейной фильтрации методом перекрытия с суммированием входной кадр расширяется последовательностью из R нулевых отсчетов. К полученной последовательности применяется алгоритм быстрого преобразования Фурье (БПФ) для перевода сигнала из временной области в частотную. Результат преобразования умножается на заранее рассчитанную частотную характеристику КИХ-фильтра (корректора). Для перевода сигнала во временную область используется алгоритм обратного быстрого преобразования Фурье (ОБПФ). Результатом преобразования является частичная свертка кадра входного сигнала $x^{(m)}[r]$ с импульсной характеристикой фильтра-корректора. Частичная свертка имеет два участка: $y_1^{(m)}[r]$ и $y_2^{(m)}[r]$. Текущий результат фильтрации, который соответствует кадру m , формируется путем суммирования:

$$y^{(m)}[r] = y_1^{(m)}[r] + y_2^{(m-1)}[r].$$

Следующим этапом обработки сигнала является компрессия динамического диапазона (КДД) с целью компенсации нарушения ощущения громкости. Главная задача КДД – это автоматическое регулирование коэффициента усиления сигнала $g(n)$ (см. рис. 3) в зависимости от уровня мощности сигнала. Характер КДД определяется функцией входа-выхода компрессора, которая показывает, как должен изменяться уровень выходного сигнала при изменении уровня входного сигнала. В работе используется компрессор широкого динамического диапазона, характеристика входа-выхода которого показана на рис. 9. Динамический диапазон входного сигнала условно разбит на три неперекрывающиеся области: «шум», «тихие звуки» и «речь». Под «шумом» понимается собственный шум внутренней схемы платформы iPhone, которая отвечает за прием, дискретизацию и квантование входного акустического сигнала. Это сигнал малой интенсивности, который при прохождении компрессора не усиливается. Выше уровня «шума» находятся «тихие звуки». Сигналы, относящиеся к этой категории, плохо различимы людьми, страдающими тугоухостью, поэтому при их обнаружении компрессор КДД начинает работать в режиме расширения, который характеризуется коэффициентом расширения

$$XR = \Delta x / \Delta y,$$

где Δx – изменение уровня входного сигнала; Δy – изменение уровня выходного сигнала.

Главной особенностью режима расширения КДД является то, что коэффициент XR всегда меньше единицы. Например, для коэффициента расширения $XR = 0,5$ изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 20 дБ.

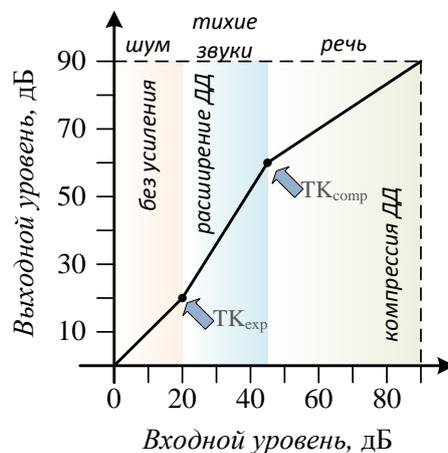


Рис. 9. Характеристика входа-выхода КДД

Если уровень входного сигнала попадает в категорию «речь», то активизируется режим компрессии динамического диапазона, характеризуемый коэффициентом компрессии

$$CR = \Delta x / \Delta y ,$$

который всегда больше единицы. Например, при коэффициенте $CR = 2$ изменение входного уровня сигнала на 10 дБ приведет к изменению выходного уровня на 5 дБ.

Таким образом, характеристика входа-выхода компрессора имеет вид кусочно-линейной кривой, имеющей характерные точки: TK_{exp} – переход к режиму расширения и TK_{comp} – переход в режим компрессии. Кроме характеристики входа-выхода компрессор имеет параметры времени срабатывания и восстановления (их смысл пояснен в разд. 2.4).

В соответствии с рис. 8 на вход компрессора поступает сигнал $y^{(m)}[r]$ от фильтра-корректора. Задача компрессора состоит в расчете линейной функции усиления $g(r)$ для формирования выходного сигнала:

$$s^{(m)}[r] = y^{(m)}[r]g[r], \quad r = 1, \dots, R.$$

Поскольку на функцию $g[r]$ наложено ограничение линейности, достаточно определить значения $g[1]$ и $g[R]$, а остальные значения находятся путем линейной интерполяции. Ниже приведен псевдокод для алгоритма определения $g[1]$ и $g[R]$:

```

P_start = 10log10(p)
for r = 1, 2, ..., R
    p_inst = (y(m)[r])2
    if (p_inst > p) then
        p = αp + (1 - α)p_inst
    else
        p = βp + (1 - β)p_inst

```

```

endif;
end for;
 $P_{end} = 10 \log_{10}(p)$ 
 $P_{start}^{out} = IO\_Func(P_{start})$ 
 $P_{end}^{out} = IO\_Func(P_{end})$ 
 $g[L] = 10^{(P_{start}^{out} - P_{start})/20}$ 
 $g[R] = 10^{(P_{end}^{out} - P_{end})/20}$ 

```

где α – параметр экспоненциального усреднения, зависящий от времени срабатывания;
 β – параметр экспоненциального усреднения, зависящий от времени восстановления.

В приведенной программе переменная p хранит текущее (среднее) значение уровня мощности сигнала. Первому кадру сигнала ($m=0$) p присваивается значение 2^{-10} , в последующем значение p сохраняется от кадра к кадру. Таким образом, P_{start} для m -го кадра равно значению P_{end} ($m-1$)-го кадра. Текущее значение мощности сигнала вычисляется путем экспоненциального усреднения. Параметр экспоненциального усреднения выбирается в зависимости от того, что происходит – нарастание или спад сигнала. Через $IO_Func()$ в программе обозначена функция входа-выхода компрессора (см. рис. 9).

Пример на рис. 10 показывает характер изменения коэффициента усиления в компрессоре в зависимости от уровня входного сигнала, а также поясняет значения понятий времени срабатывания и времени восстановления.

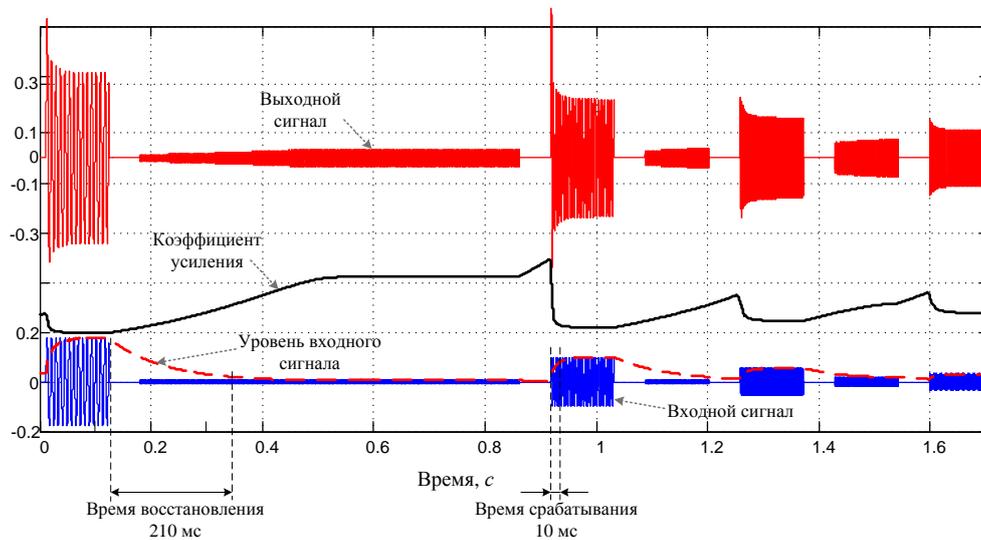


Рис. 10. Обработка тестовых тональных сигналов компрессором динамического диапазона

На рис. 11 видно, что совместное использование фильтра-корректора и компрессора динамического диапазона позволяет выполнить уровне- и частотно-зависимую обработку сигнала. Особенно следует отметить обработку фрагмента речевого сигнала, отвечающего шипящему звуку «с». Значительное усиление этого звука обусловлено тем, что его основная энергия лежит в верхней части частотного диапазона, который плохо воспринимается людьми, страдающими тугоухостью. По тем же причинам значительное усиление получил взрывной звук «п».

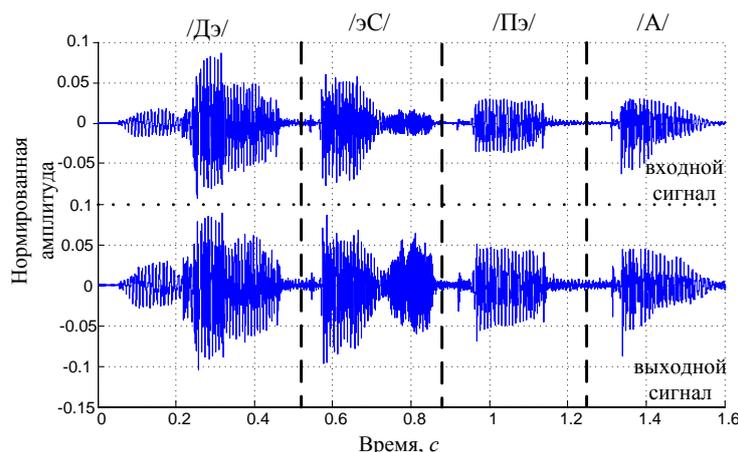


Рис. 11. Пример обработки речевого сигнала предлагаемым слуховым аппаратом на базе мобильной платформы

3.4. Задержка обработки сигнала

Задержка обработки сигнала складывается из групповой задержки корректирующего фильтра и задержки ввода-вывода аудиосистемы вычислительной платформы. Учитывая, что используется фильтр с линейной ФЧХ, групповая задержка корректирующего фильтра равна половине длительности импульсной характеристики. Для формирования фильтра с удовлетворительной аппроксимацией целевой частотной характеристики нужно использовать 128–256 коэффициентов, что при частоте дискретизации 44,1 кГц соответствует групповой задержке 1,45–2,9 мс. Организация ввода-вывода включает неустраняемые задержки ЦАП и АЦП, которые составляют примерно 2–4 мс, и задержки буфера сигнала. В отличие от СА, использующих специальные вычислители, на платформе iPhone нет возможности выполнять захват сигнала короткими фреймами в несколько отсчетов, поскольку это приводит к неоправданно высоким затратам на диспетчеризацию и обслуживание приложения со стороны операционной системы. В зависимости от модели устройства приемлемым является ввод и вывод сигнала блоками по 128–256 отсчетов, что соответствует задержке 2,9–5,8 мс. Таким образом, общая расчетная задержка обработки сигнала при реализации слухового аппарата на платформе iPhone может составлять от 6,3 до 15,7 мс.

4. Результаты экспериментов

4.1. Конфигурация устройства

Для экспериментов использовалось устройство iPhone5 со штатной гарнитурой. Операционная система iOS 7.0 позволяет применять встроенный микрофон, находящийся на панели устройства, одновременно с наушниками гарнитуры. Экспериментальным путем установлено, что встроенный микрофон обеспечивает более высокое соотношение сигнал/шум и поэтому для приложения является предпочтительным. Кроме того, большое расстояние между наушниками гарнитуры и встроенным микрофоном исключает появление АОС. Учитывая перечисленные особенности, во всех экспериментах, проведенных в рамках этой работы, использовался встроенный микрофон устройства с частотой дискретизации сигнала 44,1 кГц и разрядностью представления данных 16 бит.

В процессе аудиометрии пользователю подавались сигналы заданной частоты в следующей последовательности: 1000, 2000, 4000, 6000, 8000, 125, 250 и 500 Гц. Сигналы синтезировались с нарастающей амплитудой от минимально возможного значения до максимального на протяжении 15 с.

4.2. Полное акустическое усиление

Акустическим усилением называется разность между выходным уровнем звукового давления, создаваемого слуховым аппаратом в имитаторе уха, и входным уровнем звукового давления, измеренным в тестовой точке. Полным акустическим усилением называется акустиче-

ское усиление слухового аппарата при линейном входе-выходе и максимальном положении регуляторов слухового аппарата.

Лабораторные измерения показали, что аудиосистема iPhone5 со стандартной гарнитурой позволяет создавать полное акустическое усиление до 30 дБ (рис. 12).

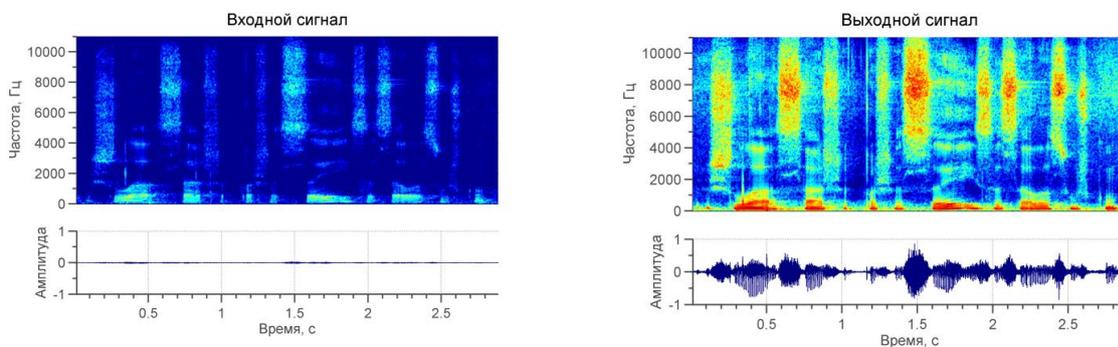


Рис. 12. Усиление речевого сигнала в слуховом аппарате

4.3. Пример обработки сигнала

Для иллюстрации работы предложенного алгоритма обработки сигнала в слуховом аппарате рассмотрим пример компенсации потери слуха. На рис. 13, а показана аудиограмма пользователя (75 лет, первая степень тугоухости). По аудиограмме выполняется расчет фильтра-корректора, в котором заложены целевые коэффициенты усиления. Для расчета АЧХ фильтра-корректора (рис. 13, б) использовалась формула Berger (см. табл. 1).

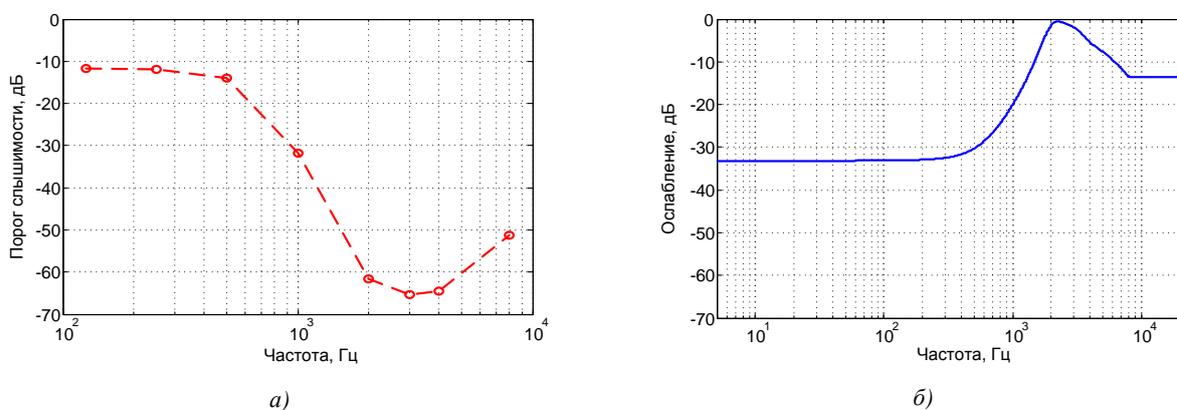


Рис. 13. Аудиометрия: а) аудиограмма пользователя; б) АЧХ фильтра-корректора

Чтобы качественно оценить результат работы предложенного алгоритма, выполнялось упрощенное моделирование потери слуха, в результате которого получался сигнал, воспринимаемый пациентом с ослабленным слухом. В общем случае система, моделирующая потерю слуха, должна отражать следующие явления, наблюдаемые при тугоухости [22]: увеличение порога слышимости, ускоренное нарастание громкости, ограничение частотной избирательности, компрессию в зависимости от уровня входного сигнала. В настоящей работе использовалась упрощенная система моделирования потери слуха, отражающая только явление увеличения порога слышимости. Для этого исходя из аудиограммы рассчитывался КИХ-фильтр методом частотной выборки.

Результат обработки речевого сигнала (рис. 14) разработанным алгоритмом коррекции слуха показан на рис. 15, а. Приведенная спектрограмма отражает существенное усиление высокочастотных составляющих речи. Вид сигнала во временной области показывает, что использование разработанного алгоритма компрессии позволяет эффективно сужать динамический диапазон сигнала, производя таким образом «выравнивание» громкости.

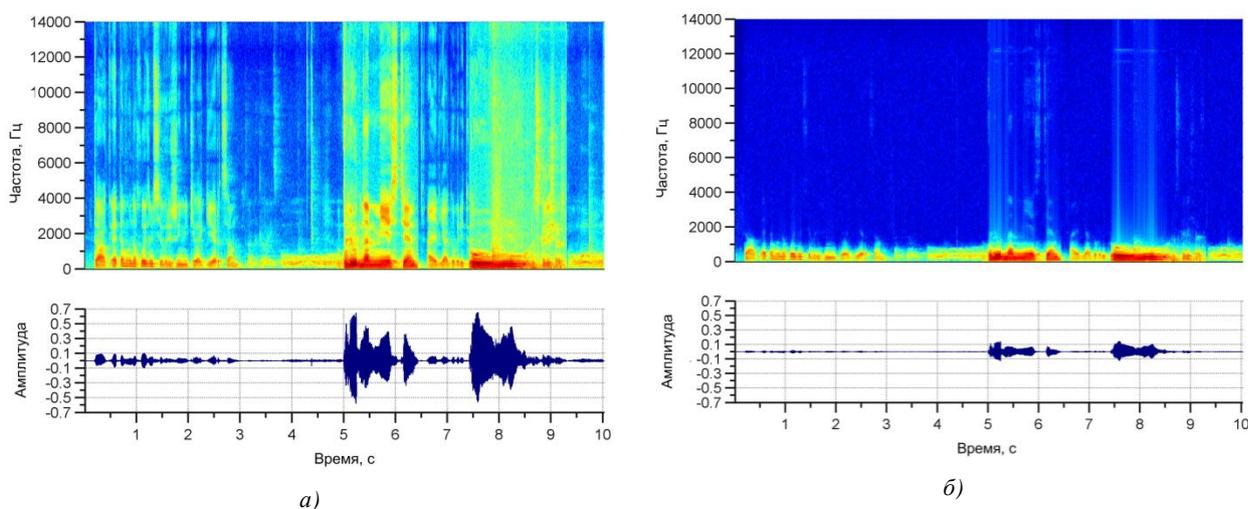


Рис. 14. Речевой сигнал: а) исходный; б) пропущенный через систему моделирования потери слуха

На рис. 15, б показан обработанный сигнал, пропущенный через систему моделирования потери слуха. Сравнение данного сигнала с изображенным на рис. 14, б выявляет превосходство первого над вторым по субъективному уровню комфортности восприятия. Обработанный сигнал лучше воспринимается за счет суженного динамического диапазона (нет слишком тихих или слишком громких звуков), а также за счет частичного сохранения высокочастотных компонентов речи, которые, как известно, вносят большой вклад в разборчивость речи.

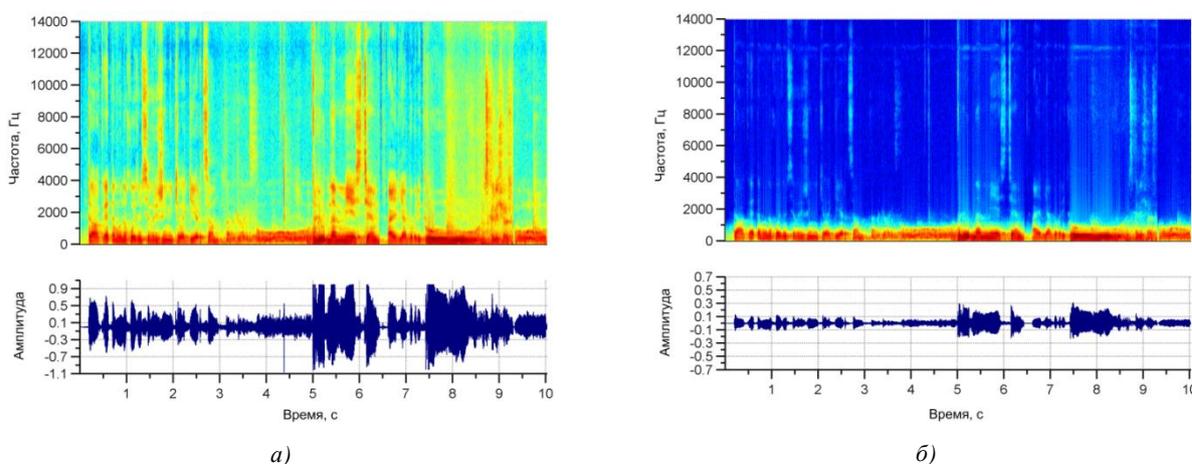


Рис. 15. Обработанный речевой сигнал: а) исходный; б) пропущенный через систему моделирования потери слуха

4.4. Разработанное приложение

Реализованная система коррекции слуха оформлена в виде специального приложения под названием Petralex (сайт приложения www.petralex.pro). Приложение позволяет сохранять настройки усиления и компрессии в виде различных пользовательских профилей, основанных на результатах тестирования слуха в каждой часто используемой акустической обстановке. Таким образом, в отличие от статических настроек специализированных СА пользователь имеет возможность быстрого переключения между профилями в зависимости от изменения окружающей акустической обстановки. На рис. 16 показаны интерфейсные экраны приложения.

Унификация операционных систем и параметров устройств позволила также реализовать возможность быстрой миграции между различными устройствами на базе операционной си-

стемы iOS с использованием облачного сервиса хранения данных iCloud. При использовании сервиса сохраненные пользовательские профили автоматически загружаются во все устройства, подключенные пользователем к персональному хранилищу данных.

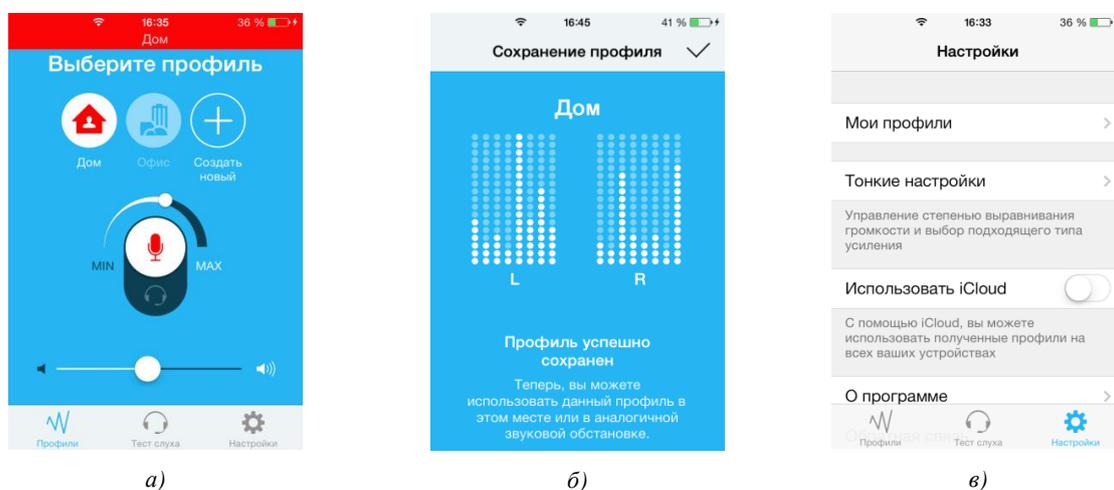


Рис. 16. Экраны приложения Petralex: а) основной интерфейсный экран приложения с регуляторами усиления; б) сохранение пользовательского профиля; в) настройка приложения

В процессе эксплуатации программы к смартфону подключается внешняя гарнитура (наушники и микрофон). Возможно подключение одних наушников. В этом случае программа использует встроенный микрофон iPhone. Пользователь может самостоятельно выбрать один из двух микрофонов при помощи переключателя, располагающегося в центре интерфейсного экрана приложения (рис. 16, а). Использование встроенного микрофона предпочтительнее, поскольку он имеет более низкий уровень собственных шумов.

Заключение

В работе показана возможность применения мобильной вычислительной платформы iPhone для реализации цифрового СА. Предложена схема обработки сигнала с малой алгоритмической задержкой, позволяющая выполнять коррекцию слуха при нейросенсорной тугоухости и учитывающая особенности вычислительной платформы. Описан алгоритм обработки, в основе которого лежат линейное частотно-зависимое усиление и широкополосная компрессия динамического диапазона сигнала. В работе приведены результаты обработки речевых сигналов с использованием предложенного алгоритма коррекции слуха, а его реализация на базе мобильного устройства позволила предложить доступную альтернативу специализированным СА.

Список литературы

1. Чекаев, Г.М. Медико-социальная характеристика заболеваний органа слуха и пути совершенствования реабилитации сурдологических больных : автореф. дис. ... канд. мед. наук / Г.М. Чекаев. – Казань, 1998. – 19 с.
2. Глухота и нарушение слуха. Информационный бюллетень № 300 Всемирной организации здравоохранения [Электронный ресурс]. – 2013. – Режим доступа : <http://www.who.int/mediacentre/factsheets/fs300/ru/index.html>. – Дата доступа : 12.11.2013.
3. Фонлантен А. Слуховые аппараты / А. Фонлантен, А. Хорст. – Ростов н/Д. : Феникс, 2009. – 304 с.
4. Pandey, A. Low-delay signal processing for digital hearing aids / A. Pandey, V.J. Mathews // IEEE Transactions on Audio, Speech, and Language Processing. – 2011. – № 4 (19). – P. 699–710.
5. Acoustic feedback reduction based on filtered-X LMS and normalized filtered-X LMS algorithms in digital hearing aids based on WOLA filterbank / R. Vincen-Bueno [et al.] // IEEE International Symposium on Intelligent Signal Processing, Alcalá, Spain, 3–5 October, 2007. – Alcalá, 2007. – P. 1–6.

6. Hellgren, J. Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification / J. Hellgren // IEEE Trans. Speech Audio Process. – 2002. – № 2 (10). – P. 119–131.
7. Farhang-Boroujeny, B. Adaptive filters: theory and applications / B. Farhang-Boroujeny. – N. Y. : Wiley, 1998. – 548 p.
8. Agnew, J. Just noticeable and objectionable group delays in digital hearing aids / J. Agnew, J.M. Thornton // J. Amer. Acad. Audiol. – 2000. – № 6 (11) – P. 330–336.
9. Stone, M.A. Tolerable hearing aids delays. II: Estimation of limits imposed during speech production / M.A. Stone, B.C.J. Moore // J. Amer. Acad. Audiol. – 2002. – № 6 (11). – P. 325–338.
10. Ryan, J. A digital signal processor for musicians and audiophiles / J. Ryan, S. Tewari // Hear. Rev. – 2009. – № 2 (16). – P. 38–41.
11. Kates, K.M. Multichannel dynamic-range compression using digital frequency warping / K.M. Kates, K.H. Arehart // EURASIP J. Appl. Signal Process. – 2005. – № 1 (18). – P. 3003–3014.
12. Harteneck, M. Design of near perfect reconstruction oversampled filter banks for subband adaptive filters / M. Harteneck, S. Weiss, R.W. Stewart // IEEE Trans. Circuits Syst. – 1999. – № 8 (46). – P. 1081–1086.
13. Biological basis of hearing-aid design / M.B. Sachs [et al.] // Ann. Biomed. Eng. – 2002. – № 2 (30). – P. 157–168.
14. Vashkevich, M. Low-delay hearing aid based on cochlear model with nonuniform subband acoustic feedback cancellation / M. Vashkevich, E. Azarov, A. Petrovsky // Proceedings of the 20th european signal proc. conf., Bucharest, Romania, 27–31 August, 2012. – Bucharest, 2012. – P. 514–518.
15. Вашкевич, М.И. Подавление эффекта акустической обратной связи в слуховых аппаратах с использованием неравнополосного банка фильтров / М.И. Вашкевич, И.С. Азаров, А.А. Петровский // Информатика. – 2012. – № 2. – С. 50–60.
16. Berger, K.W. Determining hearing aid gain / K.W. Berger, E.N. Hagberg, R.L. Rane // Hearing Instruments. – 1980. – № 30. – P. 26–44.
17. McCandless, G.A. Prescription of gain/output (POGO) for hearing aids / G.A. McCandless, P.E. Lyregaard // Hearing Instruments. – 1983. – № 34. – P. 16–21.
18. Byrne, D. The national acoustic laboratories (NAL) new procedure for selecting the gain and frequency response of a hearing aid / D. Byrne, H. Dillon // Ear and Hearing. – 1986. – № 7. – P. 257–265.
19. Ching, T. Comparison of severely and profoundly hearing-impaired children's preferences with the NAL-RP and the DSL 3.0 prescriptions // T. Ching, P. Newall, D. Wigney // Scandinavian Audiology. – 1997. – № 26. – P. 219–222.
20. Oppenheim, A.V. Digital Signal Processing / A.V. Oppenheim, R.W. Schaffer. – New Jersey : Prentice Hall, Englewood Cliffs, 1983. – 784 p.
21. Schaub, A. Digital hearing aids / A. Schaub. – N. Y. : Thieme, 2008. – 190 p.
22. Simulation of hearing loss using compressive gammachirp auditory filters / H. Hu [et al.] // International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May, 2011. – Prague, 2011. – P. 5428–5431.

Поступила 12.02.2014

¹Белорусский государственный университет
информатики и радиоэлектроники,
Минск, П. Бровки, 6
e-mail: azarov@bsuir.by,
vashkevich@bsuir.by,
palex@bsuir.by

²Республиканский научно-практический
центр отоларингологии,
Минск, ул. Сухая, 8
e-mail: kozlowasweta@tut.by

E.S. Azarov, M.I. Vashkevich, S.V. Kozlova, A.A. Petrovsky

**HEARING CORRECTION SYSTEM
BASED ON MOBILE COMPUTING PLATFORM**

An implementation of digital hearing aid based on the mobile computing platform iPhone is proposed. The developed signal processing scheme has a low delay, takes into account characteristics of the computing platform and allows to perform the correction of sensorineural hearing loss. The correction is carried out by linear frequency-dependent amplification and wideband dynamic range compression of the signal. In order to take into consideration magnitude characteristic of loudspeakers, the audiometry is performed using iPhone directly. The experimental results of with proposed hearing aid are given.

УДК 004.934.5

Ю.С. Гецэвіч, Б.М. Лабанаў, Д.А. Пакладок

ФАНЕТЫЧНАЯ І АЛАФОННАЯ АПРАЦОЎКА ТЭКСТУ Ў СІНТЭЗАТАРЫ БЕЛАРУСКАГА І РУСКАГА МАЎЛЕННЯ ДЛЯ МАБІЛЬНЫХ ПЛАТФОРМАЎ

Апісваюцца метады пераўтварэння «графема – фанема» і «фанема – алафон» для беларускай і рускай моў на аснове экспертных правіл. Прапануюцца алгарытмы пераўтварэння «графема – фанема» і «фанема – алафон» для сінтэзу маўлення па тэксце для мабільных платформаў.

Уводзіны

Фанетычны працэсар у складзе сістэмы сінтэзу маўлення па тэксце ажыццяўляе пераўтварэнне арфаграфічнага тэксту ў паслядоўнасць алафонаў. У фанетычным працэсары закладзены правілы пераўтварэння арфаграфічнага тэксту ў паслядоўнасць фанем (пераўтварэнне «графема – фанема») і правілы пераўтварэння паслядоўнасці фанем у алафонную паслядоўнасць (пераўтварэнне «фанема – алафон») [1].

На дадзены момант вылучаюць тры метады пераўтварэння «графема – фанема» [2]: пошук па слоўніку, метады на аснове экспертных правіл, метады кіруемых дадзеных.

Выкарыстанне слоўніка для запісу ўсіх фанетычных транскрыпцый словаформаў мовы дае найменшую вылічальную складанасць, аднак такі слоўнік займае шмат камп'ютарнай памяці. Асабліва гэта крытычна для флектыўных моў (да якіх належаць беларуская і руская мовы), таму што колькасць словаформаў у мове перавышае 2 млн.

Метады кіруемых дадзеных заснаваны на правілах, аднак у дадзеным выпадку гэтыя правілы выводзяцца алгарытмам на этапе навучання і аптымізуюцца для наступнай апрацоўкі з дапамогай выраўноўвання, дрэў, графаў, нейронавых сетак, статыстычных метадаў, часткова экспертам. Такім чынам здымаецца нагрузка на эксперта па вывадзе правіл і з'яўляецца магчымаць саманавучання сістэмы ў выніку папаўнення навучальнага корпуса. Дададзены метады з'яўляюцца самым дакладным пры наяўнасці, па-першае, даволі вялікага корпуса фанетычна размечаных тэкстаў і дазволу на выкарыстанне амаль неабмежаваных вылічальных рэсурсаў і памяці для прымянення ў сінтэзе маўлення атрыманай колькасці правіл. Аднак пабудова вялікага корпуса, які быў бы размечаны фанемамі і алафонамі, з'яўляецца найбольшай складанасцю для выкарыстання метаду кіруемых правіл у сінтэзе маўлення.

Пераўтварэнне на правілах з'яўляецца класічнай алгарытмізацыяй ведаў пра мову. Выкарыстанне замацаваных правіл вымаўлення дае добрыя вынікі пры досыць невялікай колькасці правіл. Аднак у гэтым метады павялічваецца вылічальная складанасць адносна першага метаду, бо ў горшым выпадку для ўваходнага слова патрэбна прымяніць кожнае правіла пераўтварэння. Таксама ўскладаецца шмат абавязкаў на экспертаў-інжынераў па распрацоўцы зразумелых, адзеленых ад алгарытмаў, проста рэдагуемых формаў запісу правіл, якія б маглі выкарыстоўваць эксперты-лінгвісты.

Практычнае выкарыстанне экспертных правіл пераўтварэння «графема – фанема» апісана для шматлікіх моў: літоўскай [3], польскай [4], грэцкай [5] і інш. Правілы пераўтварэння «графема – фанема – алафон» для рускага маўлення на цяперашні час фармалізаваныя [6] і рэалізаваныя без аддзялення ад праграмавага коду ў сінтэзатары маўлення Multiphone [1]. Таму яны не зручныя для хуткага рэдагавання і маштабавання экспертам-лінгвістам, калі знаходзяцца памылкі ў алафанізацыі слоў. Таксама гэтыя правілы з-за прывязкі да мовы праграмавання C++ не могуць быць адразу ж выкарыстаны для Java- і PHP-платформаў. Так як правілы пераўтварэння «графема – фанема – алафон» беларускага сінтэзатара маўлення былі перапрацаваны з правіл сінтэзатара рускага маўлення [7], то складанасць іх дапрацоўкі для беларускай мовы таксама прысутнічае.

На мабільныя камп'ютары накладваюцца абмежаванні на памер выкарыстання аператыўнай і пастаяннай памяці, а таксама на дазволеную колькасць працэсарных апераций у адзінку часу [8]. Таму ў дадзенай працы прапануецца метада пераўтварэння «графема – фанема – алафон», які заснаваны на экспертных правілах. Таксама адзначаецца магчымасць выкарыстання гэтага падыходу і для стварэння фанетычнага працэсара сінтэзатара маўлення для інтэрнэт-платформаў. Для фармалізацыі форм правіл і саміх правіл фанетызацыі і алафанізацыі беларускіх слоў выкарыстоўвалася адпаведная літаратура [9–13].

1. Метада пераўтварэння «графема – фанема» для дзвюх моў

Найбольш распаўсюджаны падыход у рамках гэтага метаду заключаецца ў апрацоўцы сімвальнай паслядоўнасці злева направа, і для кожнага сімвала ўваходнай паслядоўнасці графем ужываецца адно або некалькі правіл для генерацыі фанемы. Відавочна, што гэтыя правілы не могуць працаваць ізалювана. У адваротным выпадку, напрыклад для рускай мовы, сімвал /u/ будзе заўсёды прыведзены да фанемы /I/. Таму замест аднаго сімвала /u/ аналізуецца яшчэ і як мінімум левы суседні сімвал, каб можна было згенераваць фанему /Y/, калі перад /u/ стаяць /ж/ або /u/. У агульным выпадку гэтыя правілы маюць выгляд

$$A \rightarrow B / D / C$$

і азначаюць, што сімвал D будзе пераўтвораны ў A, калі левыя сімвалы – гэта B, а правыя – C.

У сістэме сінтэзу маўлення такія запісы адлюстроўваюць замацаваныя правілы вымаўлення ў мове. Заўважым, што парадак прымянення правіл з'яўляецца важным. Напрыклад, /тм/ павінна быць пераўтворана ў /Т'/, у той час як /тмсья/ ў /С,С,А/. Адпаведна другое правіла павінна выконвацца раней за першае. Такое парадкаванне правіл ускладаецца на эксперта-лінгвіста.

Беларуская і руская мовы належаць да ўсходнеславянскай групы моў. Пісьмо і чытанне вядзецца злева направа. Для іх, як і для ўсіх славянскіх моў, характэрна наяўнасць двух эфектаў асіміляцыі папярэдняй зычнай фанемы: па глухасці-звонкасці, па цвёрдасці-мяккасці. Прычым неабходна заўважыць, што эфект асіміляцыі па глухасці-звонкасці можа быць унутрыслоўным і міжслоўным. Пры гэтым іх распаўсюджванне на суседнія графемы ідзе з процілеглага чытанню боку, гэта значыць справа налева. Так як пазначаныя эфекты не ўплываюць адзін на аднаго, то можна размежаваць метада пераўтварэння «графема – фанема» на чатыры паслядоўных этапы:

1) праверка графемы на адпаведнасць правілам, якія ўлічваюць кананічныя змены, а таксама праверка на эфекты асіміляцыі зычных фанем па глухасці-звонкасці, і замена ў выпадку супадзення на адпаведную фанему або групу фанем;

2) замена літары на фанему па «стандартных» правілах;

3) праверка мяккасці папярэдняй графемы (неабходная, але недастатковая ўмова для мяккасці);

4) праверка графемы на адпаведнасць правілам змякчэння і дадаванне мяккасці да дадзенай фанемы ў выпадку супадзення.

Такім чынам, прыведзены метада можна прымяніць для апрацоўкі дзвюх моў з дапамогай адпаведных пэўнай мове экспертных правіл. Інакш кажучы, ён з'яўляецца мованезалежным, бо для апрацоўкі пэўнай мовы неабходна змяніць толькі ўваходныя рэсурсы (тэкст і фармалізаваныя правілы пераўтварэння «графема – фанема»).

Структура экспертных правіл складаецца з чатырох блокаў:

1. «Стандартных» правіл замены графемы на фанему, г. зн. найбольш частотныя замены. Напрыклад, графема «А» у беларускай мове за частую замяняецца на фанему «А».

2. Выключэнні з «стандартных» правіл замены ў выглядзе рэгулярных выказаў. Тут размяшчаюцца правілы, якія супярэчаць звычайнай замене. Гэта можа быць эфект асіміляцыі, канічная замена і г. д. Напрыклад, у рускай мове графема «И» пераходзіць у фанему «Й», калі перад ёй стаіць графема «Ж», «Ш» ці «Ц».

3. Змякчальныя графемы, г. зн. тыя графемы, перад якімі можа з'яўляцца эфект змякчэння.

4. Правілы змякчэння у выглядзе рэгулярных выказаў. Тут апісваецца неабходная умова, пры якой дадзеная графема перайдзе у мяккую фанему.

Прыклады фармата фармалізаваных правіл пераўтварэння «графема – фанема» прадстаўлены ў табл. 1.

Табліца 1

Фрагменты правіл пераўтварэння «графема – фанема» для рускай і беларускай моў

Руская мова	Беларуская мова
#Общие правила буква-фонема	#Агульныя правілы графема-фанема
А-А	А-А
Б-В	Б-В
В-В	В-В
...	...
Б-0	Э-Е
Э-Е	Ю-У
Ю-У	Я-А
Я-А	'-0
#Исключения из общих правил	#Выключэнні з агульных правілаў
(O)[+=]-O	[ГКХ][_V](I)[_V]-I
[ЖШЦ](И)-У	[АЕЁЮУЫЭЮЯ][_V](I)[_V]-J'
...	...
(Л)[Н][Ц]-0	[Г](Д)[Т]-0
#Смягчающие графемы	#Змякчальныя графемы
Е	Е
...	...
Ь	Ь
#Общие правила смягчения	#Агульныя правілы змякчэння
((XМНЛРЬПВФДТЗСГК))[ЕЁЮЯИЬ]	((XМНЛБПВФЗСГКЦ))[ЕЁЮЯИЬ]
...	...
((ГК))[ГКХ]	(С)[Ш]

2. Алгарытм пераўтварэння «графема – фанема»

Алгарытм пераўтварэння «графема – фанема» замяняе ўсе літары сінтагмы на іх фанетычныя аналагі і складаецца з сямі этапаў:

- 1) падрыхтоўкі дадзеных (крокі 1–4);
- 2) апрацоўкі службовых сімвалаў (крокі 5, 8);
- 3) фанемнага пераўтварэння па выключэнням з правіл (крок 6);
- 4) фанемнага пераўтварэння па «стандартным» правілам (крок 7);
- 5) праверкі мяккасці папярэдняй графемы (крок 9);
- 6) апрацоўкі змякчэння фанемы па правілам змякчэння (крок 10);
- 7) апрацоўкі вынікаў алгарытма (крокі 11, 12).

Уваходныя дадзеныя алгарытму:

Тэкст з прастаўленымі пазіцыямі націскаў і інтанацыйнымі меткамі S.

Рэсурсы алгарытму:

1. Мноства галосных фанем $Vow = \{Vow_1, \dots, Vow_{NVow}\}$, дзе $NVow$ – колькасць галосных фанем; мноства абазначэнняў раздзяляльнікаў $D = \{D_1, \dots, D_{ND}\}$ (заўважым, што D_1 з'яўляецца сімвалам пачатку або канца сінтагмы), дзе ND – колькасць абазначэнняў раздзяляльнікаў; мноства абазначэнняў націскаў $Sstr = \{Sstr_1, \dots, Sstr_{NSstr}\}$, дзе $NSstr$ – колькасць абазначэнняў націскаў; службовы сімвал змякчэння $Soft = /' /$.

2. Правілы «графема – фанема» $HV = \{ \langle H_1, V_1 \rangle, \dots, \langle H_{NHV}, V_{NHV} \rangle \}$, дзе NHV – колькасць сімвалаў мовы $H = \{H_1, \dots, H_{NHV}\}$ і адпаведных фанем у мове $V = \{V_1, \dots, V_{NHV}\}$; выключэнні з правіл «графема – фанема» $RF = \{ \langle R_1, F_1 \rangle, \dots, \langle R_{NRF}, F_{NRF} \rangle \}$, дзе NRF – колькасць пар правіл

выключэнняў у выглядзе рэгулярных выразаў $R=\{R_1, \dots, R_{NRF}\}$ і адпаведныя ім фанемы $F=\{F_1, \dots, F_{NRF}\}$; спіс змякчальных фанем $M=\{M_1, \dots, M_{NM}\}$, дзе NM – колькасць змякчальных фанем; правілы змякчэння ў выглядзе рэгулярных выразаў $E=\{E_1, \dots, E_{NE}\}$, дзе NE – колькасць правіл змякчэння суседняй фанемы.

Выхадныя дадзеныя:

1. Паслядоўнасць фанетычных сінтагмаў $Sph = \bigcup_{i=1}^{NSph} Sph_i$, $Sph_i = \bigcup_{j=1}^{NPh} Ph_j$,
 $i = 1, \dots, NSph$.

2. Паслядоўнасць колькасці складоў у кожнай фанетычнай сінтагме $NSyl = \bigcup_{i=1}^{NSph} Nsyl_i$.

Спецыяльныя функцыі:

$Regex(t, r)$ – функцыя, якая ажыццяўляе пошук рэгулярнага выразу r у тэкставым радку t .

$$Regex(t, r) = \begin{cases} 0, & \text{калі } t \text{ не адпавядае рэгулярнаму выразу } r; \\ 1, & \text{калі } t \text{ адпавядае рэгулярнаму выразу } r. \end{cases}$$

Алгарытм:

Уваход:

Крок 1. Лічым, што $S = \bigcup_{i=1}^{NS} S_i$, дзе NS – колькасць сінтагмаў уваходнага тэксту.

Крок 2. Для кожнай сінтагмы S_i , $i = 1, \dots, NS$ выконваем крокі 3–12. Далей *крок 13*.

Крок 3. Лічым індикатар мяккасці роўным нулю, гэта значыць $Flag:=0$.

Крок 4. Для кожнай паслядоўнасці сімвалаў $X_{j-2}, X_{j-1}, X_j, X_{j+1}, X_{j+2}$, $j = NX, \dots, 1$ з сінтагмы S_i ($S_i = \bigcup_{j=1}^{NX} X_j$, NX – колькасць сімвалаў у сінтагме) выконваем *крокі 5–11*. Прычым павінна выконвацца сістэма ўмоў:

$$\begin{cases} X_{j-2} := NULL, j-2 < 1; \\ X_{j-1} := NULL, j-1 < 1; \\ X_{j+1} := NULL, j+1 > NX; \\ X_{j+2} := NULL, j+2 > NX. \end{cases}$$

Калі сімвалы ў дадзенай сінтагме S_i скончыліся, пераходзім да *кроку 2*.

Крок 5. Калі сімвал X_j з'яўляецца раздзяляльным ці абазначэннем націску, г. зн. $X_{j+1} \in (D \cup Sstr)$, то лічым $TEMP:=X_j$ і пераходзім да *кроку 11*.

Крок 6. Для кожнага рэгулярнага выразу R_k , $k = \{1, \dots, NRF\}$ вылічыць $Regex(X_{j-2} \cup X_{j-1} \cup X_j \cup X_{j+1} \cup X_{j+2}, R_k)$. Калі для некаторага k умова $Regex(X_{j-2} \cup X_{j-1} \cup X_j \cup X_{j+1} \cup X_{j+2}, R_k) \neq 0$ выконваецца, то лічым $TEMP:=F_k$, і далей *крок 8*.

Крок 7. Для кожнага H_m , $m = 1, \dots, NHV$ правяраем, ці супаў сімвал мовы H_m з наяўным сімвалам X_j . Калі для некаторага m умова $H_m = X_j$ была выканана, то знаходзім адпаведную ёй фанему з пары $\langle H_m, V_m \rangle$ і лічым $TEMP:=V_m$.

Крок 8. Калі бягучая фанема будзе роўная пустому мноству $TEMP=NULL$, то перайсці да *кроку 4*.

Крок 9. Калі сімвал X_{j+1} змякчальны, г. зн. $X_{j+1} \in M$ або $Flag:=1$, то перайсці да *кроку 10*. Інакш лічым $Flag = 0$ і пераходзім да *кроку 11*.

Крок 10. Для кожнага правіла змякчэння E_n , $n = 1, \dots, NE$, вылічваем $Regex(X_j \cup X_{j+1}, E_n) \neq 0$ выконваецца,

то дадаем сімвал змякчэння да фанемы $TEMP$, г. зн. лічым $TEMP := TEMP \cup Soft$, таксама мяняем індыкатар мяккасці $Flag:=1$.

Крок 11. Дадаванне $TEMP$ да выніковай паслядоўнасці фанем сінтагмы Sph_i знойдзенай фанемы $TEMP$, а менавіта $Sph_i := TEMP \cup Sph_i$. Калі $TEMP \in Vow$, то $Nsyl_i := Nsyl_i + 1$, дзе $Nsyl_i$ – колькасць складоў у сінтагме Sph_i .

Крок 12. Дадаванне фанетычнай сінтагмы Sph_i да выхаднога фанетычнага тэксту Sph , г. зн. $Sph := Sph \cup Sph_i$. Захаванне колькасці складоў у сінтагме Sph_i у паслядоўнасці $Nsyl$, г. зн. $Nsyl := Nsyl \cup Nsyl_i$, што абазначае колькасць складоў адпаведна кожнай раней апрацаванай сінтагме.

Крок 13. Канец алгарытму.

3. Метад пераўтварэння «фанема – алафон» для дзвюх моў

Дадзенае пераўтварэнне ажыццяўляе замену фанемы на кадыфікаваную пазнаку алафона. Алафон – натуральны гук, які абумоўлены пэўным фанетычным асяроддзем. У адрозненне ад фанемы алафон з'яўляецца не абстрактным паняццем, а пэўным маўленчым гукам. У акустычным працэсары кадыфікаваная пазнака алафона будзе ўказваць на адпаведны алафон з акустычнай базы, які будзе мадыфікавацца і склейвацца з іншымі алафонамі ў адзін алафонны радок – маўленчы сігнал. Для спрошчанага апісання працэсаў утварэння гуку ў артыкуле будзем ўсюды выкарыстоўваць паняцце алафон замест выразу *кадыфікаваная пазнака алафона*.

У якасці рэалізацыі агульнага алгарытму пераўтварэння «фанема – алафон» для мабільных прыстасаванняў была задзейнічана спрошчаная экспертная сістэма: правілы запісаны ў выглядзе класіфікацыі груп цэнтральнага элемента, левага і правага кантэксту, а таксама выхаднога значэння цэнтральных груп фанем для ўсіх магчымых варыяцый кантэкстаў. Для дасягнення аптымальных суадносін прадукцыйнасці і спажывання рэсурсаў экспертных правілы пераўтвараюцца сістэмай у машынныя хэш-коды падчас ініцыялізацыі сінтэзу маўлення. Такая схема дазваляе пазбегнуць разбору і захоўвання значэнняў для кожнага элемента пры ўсіх магчымых варыяцыях кантэкстаў пры невялікім павелічэнні вылічальнай складанасці алгарытму, г. зн. памяншэнне выкарыстання аператыўнай памяці на 70 % пры павелічэнні на 5–10 дадатковых ітэрацый алгарытма пошуку адпаведнага кантэксту для алафона.

Правілы пераўтварэння «фанема – алафон» уяўляюць сабой вынік эксперыментальнага даследавання розных фанетычных груп па іх фанетычным асяроддзі. Пасля аналізу тэарэтычных выкладак [1, 3] і практычнай рэалізацыі літарафанамаалафоннага працэсара сістэмы сінтэзу маўлення Multiphone быў распрацаваны зручны фармат запісу правіл пераўтварэнняў «фанема – алафон». Характэрна, што такі фармат правіл дазваляе адначасова запісваць і выкарыстоўваць правілы пераўтварэння як для беларускай, так і для рускай мовы (табл. 2).

Структура экспертных правіл складаецца з наступных блокаў:

- 1) алфавіту сімвалаў, якія могуць утвараць алафоны. Уключае ў сябе фанемы мовы і сімвал пачатку ці канца сінтагмы;
- 2) цэнтральных груп фанем (фанетычных груп, характэрных для пэўнай цэнтральнай фанемы);
- 3) груп фанем левага кантэксту, якія могуць уплываць на фанетычнае асяроддзе пэўнай цэнтральнай фанемы злева;
- 4) груп фанем правага кантэксту, якія могуць уплываць на фанетычнае асяроддзе пэўнай цэнтральнай фанемы зправа;
- 5) правілаў пераўтварэння для кожнай цэнтральнай групы фанем, якія суадносяць для цэнтральнай групы фанем пэўныя значэнні левага і правага кантэкстаў алафонаў.

Правілы пераўтварэння «фанема – алафон» для дзвюх моў (фрагменты)

Назва групы	Значэнні (скарочана)
Алфавіт сімвалаў	A0,A1,A2,A3,E0,E1,E2,E3,I0,I1,I2,I3,O0,O1,O2,O3,Y0,Y1,Y2,Y3,U0,U1,U2,U3,B,B',D,D',G,G',Z,Z',ZH,L,L',M,M',N,N',P,P',T,T',K,K',C,CH',F,F',S,S',H,H',SH,SH',V,V',J',R,R',C',CH,SCH,GH,DZ',DZH,GH',W,#,
Цэнтральныя групы фанем	C0: A0,A1,A2,A3,E0,E1,E2,E3,I0,I1,I2,I3,Y0,Y1,Y2,Y3,
	C1: U0,U1,U2,U3,O0,O1,O2,O3,
	...
Групы фанем левага кантэксту	C6: R,R',
	L0: #,
	L1: P,B,F,V,M,U0,U1,U2,U3,O0,O1,O2,O3,W,
	L2: SH,ZH,R,T,C,S,D,Z,N,L,A0,A1,A2,A3,E0,E1,E2,E3,Y0,Y1,Y2,Y3,CH,SCH,DZH,
Групы фанем правага кантэксту	...
	L11: B,B',D,D',G,G',Z,Z',ZH,L,L',M,M',N,N',P,P',T,T',K,K',C,CH',F,F',S,S',SH,SH',H,H',V,V',J',R,R',C',CH,SCH,GH,DZ',DZH,GH',W,#,
	R0: #,
	R1: P,B,F,V,M,L,W,O0,O1,O2,O3,U0,U1,U2,U3,
Правілы пераўтварэння для кожнай цэнтральнай групы фанем	...
	R14: A0,A1,A2,A3,E0,E1,E2,E3,Y0,Y1,Y2,Y3,I0,I1,I2,I3,O0,O1,O2,O3,U0,U1,U2,U3,
	C0
	L0R0-00
	L5R0-10
	L6R0-20
	...
	C1
	L0R0-00
	L1R0-10
L2R0-20	
...	

4. Алгарытм пераўтварэння «фанема – алафон»

Алгарытм пераўтварэння «фанема – алафон» складаецца з сямі этапаў:

- 1) падрыхтоўкі дадзеных (крокі 1–4);
- 2) вызначэння першага індэкса алафона (крокі 5–11);
- 3) апрацоўкі службовых сімвалаў (крок 12);
- 4) вызначэння мноства фанемных груп для цэнтральнай, правай і левай фанем (крок 13);
- 5) пошука такога варыянта сярод правіл для вызначанай фанемнай групы цэнтральнай фанемы, пры якім існуе пэўнае значэнне левага і правага кантэкстаў алафона – другі і трэці індэкс алафона (крокі 14–16);
- 6) знаходжання поўнага значэння алафона (крок 17);
- 7) апрацоўкі вынікаў алгарытма (крокі 18, 19).

Уваходныя дадзеныя алгарытма:

1. Паслядоўнасць фанетычных транскрыпцый сінтагмаў уваходнага тэксту *Sph*.
2. Паслядоўнасць колькасцяў складоў у кожнай сінтагме *Nsyl*.

Рэсурсы алгарытма:

1. Мноства галосных фанем $Vow = \{Vow_1, \dots, Vow_{NVow}\}$, дзе $NVow$ – колькасць галосных фанем; мноства зычных фанем $Con = \{Con_1, \dots, Con_{NCon}\}$, дзе $NCon$ – колькасць зычных фанем; мноства раздзяляльнікаў $D = \{D_1, \dots, D_{ND}\}$ (заўважым, што D_1 з'яўляецца знакам пачатку або

канца сінтагмы), дзе ND – колькасць раздзяляльнікаў; мноства абазначэнняў націскаў $Sstr = \{Sstr_1, \dots, Sstr_{NSstr}\}$, дзе $NSstr$ – колькасць абазначэнняў націскаў, $Sstr_1$ – сімвал поўнага націску, $Sstr_2$ – сімвал частковага націску.

2. Фанемныя групы для цэнтральных фанем $C = \{<NameC_1, C_1>, \dots, <NameC_{NC}, C_{NC}>\}$, $m = 1, \dots, NC$, дзе $C_m = \{PhC_{m,r} \in \{Vow \cup Con\} | r = 1, \dots, NC_m\}$, C_m – мноства фанем групы, $NameC_m$ – унікальная назва групы. Заўважым, што фанемы $PhC_{m,r}$ выкарыстоўваюцца ў C не больш аднаго разу.

Фанемныя групы для левых фанем $L = \{<NameL_1, L_1>, \dots, <NameL_n, L_n>\}$, $n = 1, \dots, NL$, дзе $L_n = \{PhL_{n,s} \in \{Vow \cup Con \cup D\} | s = 1, \dots, NL_n\}$, L_n – мноства фанем групы, $NameL_n$ – унікальная назва групы.

Фанемныя групы для правых фанем $R = \{<NameR_1, R_1>, \dots, <NameR_{NR}, R_{NR}>\}$, $p = 1, \dots, NR$, дзе $R_p = \{PhR_{p,t} \in \{Vow \cup Con \cup D\} | t = 1, \dots, NR_p\}$, R_p – мноства фанем групы, $NameR_p$ – унікальная назва групы.

Заўважым, што кожная група фанем мае ўнікальную назву, напрыклад «C2», «L1», «R13». Першы элемент назвы адлюстроўвае прыналежнасць да месца размяшчэння фанемы (C – групы для цэнтральных фанем, L – групы для левых фанем, R – групы для правых фанем), другі вызначае склад фанем, які ўваходзіць у дадзеную групу.

3. Для кожнай фанемнай групы C_m , $m = 1, \dots, NC$, вызначаны правілы пераўтварэння «фанема – алафон» $RA_m = \{<LR_1, Index_1>, \dots, <LR_{NRam}, Index_{NRam}>\}$, $NRam$ – колькасць пар элементаў правіл $LR = \{LR_1, \dots, LR_{NRam}\}$ і $Index = \{Index_1, \dots, Index_{NRam}\}$.

Выхадныя дадзеныя:

1. Паслядоўнасць алафонных транскрыпцый сінтагмаў Sal .

Функцыя $F(LR_i, NameL_j, NameR_k)$ правярае, ці роўны першы элемент правіл пераўтварэння «фанема – алафон» LR_i аб'яднанню назваў груп для левых $NameL_j$ і правых $NameR_k$ фанем:

$$F(LR_i, NameL_j, NameR_k) = \begin{cases} 0, & LR_i \neq NameL_j \cup NameR_k; \\ 1, & LR_i = NameL_j \cup NameR_k. \end{cases}$$

Алгарытм:

Крок 1. Лічым, што $Sph = \bigcup_{i=1}^{NSph} Sph_i$, $Nsyl = \bigcup_{i=1}^{NSph} Nsyl_i$, $NSph$ – колькасць сінтагмаў з распісанымі фанетычнымі транскрыпцыямі.

Крок 2. Для кожнай сінтагмы ў фанемным выглядзе Sph_i , $i = 1, \dots, NSph$ (заўважым, што $Sph_i = \bigcup_{j=1}^{NPh} Ph_j$, NPh – колькасць фанем у сінтагме) выконваем крокі 3–19. Далей крок 20.

Крок 3. Лічым наяўную колькасць складоў у сінтагме Sph_i роўнай $NsylT:=0$;

Крок 4. Для кожнай паслядоўнасці фанем Ph_{j-1} , Ph_j , Ph_{j+1} сінтагмы Sph_i выконваем крокі 5–18, прычым $j = NPh, \dots, 1$. Пры гэтым павінна выконвацца сістэма ўмоў:

$$\begin{cases} Ph_{j-1} := D_1, & j-1 < 1; \\ Ph_{j+1} := D_1, & j+1 > NPh. \end{cases}$$

Калі фанемы ў дадзенай сінтагме скончыліся, пераходзім да кроку 2.

Крок 5. Калі фанема $Ph_j \in Vow$, то пераходзім да кроку 6. Інакш пераходзім да кроку 10.

Крок 6. Павялічваем колькасць складоў у сінтагме Sph_i на адзінку $NsylT:=NsylT+1$.

Крок 7. Калі $Ph_{j+1} = Sstr_1$ ($Sstr_1 = /+ /$), то дадзеная фанема мае поўны націск, тады першы індэкс алафона $IndexA$ прымае значэнне 0, а менавіта $IndexA:=0$. Тады за фанему Ph_{j+1} прымаем фанему з індэксам $j+2$ з паслядоўнасці фанем Sph_i , а ў паслядоўнасці фанем Sph_i выдаляем фанему з індэксам $j+1$, агульную колькасць фанем памяншаем на адзінку ў паслядоўнасці фанем Sph_i $NPh:=NPh-1$. Запамінаем, што наяўная галосная Ph_j – націскная праз $FlagSV:=1$, і пераходзім да кроку 13.

Крок 8. Калі $Ph_{j+1} = Sstr_2$ ($Sstr_2 = \neq$), то дадзеная фанема мае частковы націск, тады першы індэкс алафона $IndexA$ прымае значэнне 1, а менавіта $IndexA:=1$. Тады за фанему Ph_{j+1} прымаем фанему з індэксам $j+2$ з паслядоўнасці фанем Sph_i , а ў паслядоўнасці фанем Sph_i выдаляем фанему з індэксам $j+1$, агульную колькасць фанем памяншаем на адзінку ў паслядоўнасці фанем $Sph_i NPh:=NPh-1$. Запамінаем, што наяўная галосная Ph_j – націскная праз $FlagSV:=1$, і пераходзім да кроку 13.

Крок 9. Калі $FlagSV = 0$ і дадзеная фанема Ph_j не знаходзіцца ў крайнім левым $NsylT \neq 1$ ці крайнім правым $NsylT \neq Nsyl_i$ складзе, то дадзеная фанема Ph_j з'яўляецца ненаціскай галоснай другой ступені рэдукцыі, тады першы індэкс алафона $IndexA$ абазначаем лічбай 3, а менавіта $IndexA:=3$. Калі інакш, яна з'яўляецца ненаціскай галоснай першай ступені рэдукцыі, тады першы індэкс алафона $IndexA$ абазначаем лічбай 2, а менавіта $IndexA:=2$. Пераходзім да кроку 13.

Крок 10. Калі дадзеная фанема зычная, г. зн. $Ph_j \in Con$, то пераходзім да кроку 11. Інакш пераходзім да кроку 12.

Крок 11. Калі фанема Ph_j роўная першай левай фанеме Ph_{j-1} адносна яе ($Ph_j = Ph_{j-1}$), то яна лічыцца падвоенай, г. зн. $IndexA:=1$. Тады за фанему Ph_{j-1} прымаем фанему з індэксам $j-2$ з паслядоўнасці фанем Sph_i , а ў паслядоўнасці фанем Sph_i выдаляем фанему з індэксам $j-1$, агульную нумарацыю наяўнай фанемы памяншаем на 1, г.зн. $j:=j-1$, бо агульная колькасць фанем паменшылася на адзінку ў паслядоўнасці фанем $Sph_i NPh:=NPh-1$.

Калі інакш, дадзеную фанему лічым адзінарнай $IndexA:=0$.

Пераходзім да кроку 13.

Крок 12. Калі Ph_j з'яўляецца раздзяляльным элементам $Ph_j \in D$, то запамінаем яго ў часовай зменнай $TEMP:=Ph_j$ і пераходзім да кроку 18.

Крок 13. Вызначаем мноства фанемных груп, у якія ўваходзіць цэнтральны элемент Ph_j :

$$Cx = \{ \langle NameC_m, C_m \rangle | PhC_{m,r} \in C_m, C_m \in C, m = 1, \dots, NC, r = 1, \dots, NC_m, \text{ калі } PhC_{m,r} = Ph_j \}.$$

Улічваючы фармат правіл пераўтварэння «графема – фанема», мноства Cx складаецца з аднаго элемента. Запамінаем індэкс m , пры якім была выканана ўмова $PhC_{m,r} = Ph_j$, г. зн. $iCx:=m$. Па індэксе iCx знаходзім адпаведныя гэтаму мноству правілы пераўтварэння RA_{iCx} .

Вызначаем мноства фанемных груп, у якія ўваходзіць левы элемент Ph_{j-1} :

$$Lx = \{ \langle NameL_n, L_n \rangle | PhL_{n,s} \in L_n, L_n \in L, n = 1, \dots, NL, s = 1, \dots, NL_n, \text{ пры ўмове } PhL_{n,s} = Ph_j \}.$$

Вызначаем мноства фанемных груп, у якія ўваходзіць правы элемент Ph_{j+1} :

$$Rx = \{ \langle NameR_p, R_p \rangle | PhR_{p,t} \in R_p, R_p \in R, p = 1, \dots, NR, t = 1, \dots, NR_p, \text{ пры ўмове } PhR_{p,t} = Ph_j \}.$$

Крок 14. Для кожнага мноства фанем знойдзенага левага кантэксту Lx_{kl} , $Lx = \{Lx_1, \dots, Lx_{NLx}\}$, $kl = 1, \dots, NLx$, дзе NLx – колькасць груп для левага кантэксту Ph_{j-1} , выконваем крокі 15–16.

Крок 15. Для кожнага мноства фанем правага кантэксту Rx_{k2} , $Rx = \{Rx_1, \dots, Rx_{NRx}\}$, $k2 = 1, \dots, NRx$, дзе NRx – колькасць груп для правага кантэксту Ph_{j+1} , выконваем крок 16.

Крок 16. Для першага элемента LR_q кожнага правіла RA_{iCx} , $q = 1, \dots, NRA_{iCx}$ правяраем значэнне функцыі $F(LR_q, NameL_{kl}, NameR_{k2})$. Калі $F(LR_q, NameL_{kl}, NameR_{k2}) = 1$, то індэкс q вызначае адпаведны другі элемент правіла RA_{iCx} – $Index_q$. Гэты элемент $Index_q$ з'яўляецца другім і трэцім індэксам і алафона $IndexB:=Index_q$.

Крок 17. Знаходзім поўнае значэнне алафона $TEMP := Ph_j \cup IndexA \cup IndexB$.

Крок 18. Дадаём $TEMP$ да алафоннай транскрыпцыі сінтагмы $Sal_i := TEMP \cup Sal_i$ і пераходзім да кроку 4.

Крок 19. Дадаём алафонную транскрыпцыю сінтагмы да выхаднага алафоннага тэксту $Sal := Sal \cup Sal_i$ і пераходзім да кроку 2.

Крок 20. Канец алгарытму.

5. Вынікі працы алгарытмаў «графема – фанема» і «фанема – алафон» для беларускай і рускай моў

Прывядзём некаторыя прыклады пераўтварэння «графема – фанема» і «фанема – алафон» для беларускай і рускай моў (табл. 3, 4).

Табліца 3

Прыклады літарафанемаалафоннага пераўтварэння слоў для беларускай мовы

Словы	Пераўтварэнне «графема – фанема»	Пераўтварэнне «фанема – алафон»
бюльбю+левы	B',U,L',B',U,+,L',E,V,Y	B'004,U243,L'001,B'002,U043,L'004,E341,V012,Y210
вэ+ндзіць	V,E,+,N',DZ',I,C'	V001,E013,N'001,DZ'004,I243,C'000
удзю+х	U,DZ',V',U,+,H	U203,DZ'001,V'001,U042,H000
льві+ца	L',V',I,+,C,A	L'001,V'001,I042,C002,A220
міжго+р'е	M',I,ZH,GH,O,+,R,J',E	M'004,I242,ZH001,GH001,O032,R001,J'002,E240
геадэ+зія	GH',E,A,D,E,+,Z',I,J',A	GH'004,E242,A222,D002,E023,Z'004,I343,J'012,A240
судзья+	S,U,DZ',DZ',A,+	S002,U223,DZ'102,A040
джу+нглі	DZH,U,+,N,GH,L',I	DZH002,U022,N001,GH004,L'004,I240
еўразо+на	J',E,W,R,A,Z,O,+,N,A	J'002,E241,W013,R012,A222,Z002,O022,N004,A220
ззя+нне	Z',Z',A,+,N',E	Z'102,A043,N'104,E240
касне+шся	K,A,S',N',E,+,S',S',A	K004,A233,S'002,N'002,E043,S'102,A240
i=ншакраі+нец	J',I,=,N,SH,A,K,R,A,J',I,+,N',E,C	J'002,I142,N003,SH002,A322,K004,R022,A223,J'011,I043,N'004,E242,C000

Табліца 4

Прыклады літарафанемаалафоннага пераўтварэння слоў для рускай мовы

Словы	Пераўтварэнне «графема – фанема»	Пераўтварэнне «фанема – алафон»
мужичо+чек	M,U,ZH,Y,CH',O,+,CH',E,K	M004,U212,ZH004,Y223,CH'001,O043,CH'002,E242,K000
кири+ллища	K',I,R',I,+,L',L',I,C,A	K'002,I243,R'002,I043,L'104,I342,C002,A220
объе+зд	A,B,J',E,+,S,T	A201,B001,J'001,E042,S002,T000
бе+лого	B',E,+,L,A,V,A	B'002,E041,L004,A311,V012,A210
проезжа+тсья	P,R,A,J',E,ZH,A,+,C,C,A	P002,R012,A223,J'012,E242,ZH102,A022,C102,A220
со+лнце	S,O,+,N,C,E	S001,O022,N003,C002,E220
расчи+тывать	R,A,SH',I,+,T,Y,V,A,T'	R022,A223,SH'001,I042,T002,Y321,V012,A213,T'000
безотчѐ+тен	B',E,Z,A,CH',CH',O,+,T',E,N	B'004,E242,Z004,A223,CH'101,O043,T'002,E242,N000
разбе+жка	R,A,Z,B',E,+,SH,K,A	R022,A222,Z001,B'002,E042,SH002,K004,A230
ию+льский	I,J',U,+,L',S,K',I,J'	I203,J'011,U043,L'003,S002,K'002,I243,J'010

Пасля таго як алгарытмы былі рэалізаваны праграма ў мабільным сінтэзатары маўлення, яны былі пратэставаны на корпусе слоў (табл. 5). Корпус слоў быў складзены паводле выбаркі ўсіх магчымых пяці сімвальных камбінацый сімвалаў у словах электронных слоўнікаў [14]. Спачатку гэты корпус быў апрацаваны стацыянарнай версіяй сінтэзатара маўлення Multiphone, а пасля – мабільнай версіяй. За эталон правільнасці быў прыняты вынік працы стацыянарнага сінтэзатара маўлення. Адносна яго была падлічана колькасць адпаведна для беларускай і рускай моў правільна апрацаваных мабільнай версіяй сінтэзатараў слоў – 82010 і 83990, фанем – 918440 і 888950, алафонаў – 903216 і 871989.

Табліца 5

Вынікі тэставання алгарытмаў для беларускай і рускай моў

Параметры	Беларуская мова	Руская мова
Колькасць слоў	102 169	101 958
Колькасць правільных слоў	82 010	83 990
Працэнтныя адносіны карэктна пераўтвораных слоў	80,269 %	82,377 %
Колькасць фанем	934 583	897 635
Колькасць правільных фанем	918 440	888 950
Працэнтныя адносіны карэктна пераўтвораных фанем	98,273 %	99,032 %
Колькасць алафонаў	934 583	897 635
Колькасць правільных алафонаў	903 216	871 989
Працэнтныя адносіны карэктна пераўтвораных алафонаў	96,644 %	97,143 %

Вынікі тэставання хуткасці алгарытмаў для беларускай і рускай моў на адабраных экспертам-лінгвістам карпусах слоў аб'ёмам каля 1000 слоў прадстаўлены ў табл. 6. У корпусе былі прадстаўлены найбольш часта ўжывальныя словы для кожнай з моў. Як бачна з табліцы, хуткасць алгарытма «графема – фанема» вельмі залежыць ад фанетычнага склада слова. Алгарытм «фанема – алафон», наадварот, не залежыць ад слова і працуе вельмі хутка.

Табліца 6

Вынікі тэставання хуткасці алгарытмаў для беларускай і рускай моў

Алгарытм	Характарыстыка	Беларуская мова	Руская мова
«Графема – фанема»	Найменшы час, мс	15	15
	Найбольшы час, мс	80	72
	Сярэдні час, мс	35	33
«Фанема – алафон»	Найменшы час, мс	1	1
	Найбольшы час, мс	2	2
	Сярэдні час, мс	1	1
«Графема – фанема – алафонная апрацоўка»	Найменшы час, мс	16	16
	Найбольшы час, мс	81	73
	Сярэдні час, мс	36	34

Заклучэнне

У артыкуле былі прапанаваны метады пераўтварэнняў «графема – фанема» і «фанема – алафон», якія заснаваны на экспертных правілах, для сінтэзу беларускага і рускага маўленняў. Прыведзеныя алгарытмы правяраюцца адносна эталоннай стацыянарнай сістэмы сінтэзу маўлення Мультыфон. Вынікі тэстаў сведчаць пра высокі ўзровень дакладнасці працы алгарытмаў для беларускай і рускай мовы пры апрацоўцы слоў (больш 80 %), фанем (больш 98 %), алафонаў (больш 96 %). Далейшая праца будзе накіравана на дапрацоўку практычнай рэалізацыі прыведзеных алгарытмаў.

Спіс літаратуры

1. Лобанов, Б.М. Компьютерный синтез и клонирование речи / Б.М. Лобанов, Л.И. Цирульник. – Минск : Беларус. навука, 2008. – 344 с.
2. Taylor, P. Text-to-Speech Synthesis / P. Taylor. – N.Y. : Cambridge University Press, 2009. – 626 p.
3. Norkevičius, G. Knowledge-based grapheme-to-phoneme conversion of Lithuanian words / G. Norkevičius, G. Raškinis, A. Kazlauskienė // SPECOM 2005, 10th Intern. Conf. Speech and Computer. – 2005. – P. 235-238.
4. Steffen-Batóg, M. An algorithm for phonetic transcription of orthographic texts in Polish / M. Steffen-Batóg, P. Nowakowski // Studia Phonetica Posnaniensia / eds. M. Steffen-Batóg, W. Awedyk. – Poznań : Wydawnictwo Naukowe UAM, 1993. – Vol. 3. – P. 135–183.

5. Chalamandaris, A. Rule-based grapheme-to-phoneme method for the Greek / A. Chalamandaris, S. Raptis, P. Tsiakoulis // 9th European Conference on Speech Communication and Technology. – 2005. – P. 2937–2940.
6. Цирульник, Л.И. Алгоритм генерации фонемной последовательности по орфографическому тексту в системе синтеза речи / Л.И. Цирульник // Информатика. – 2006. – № 4. – С. 61–70.
7. Гецевич, Ю.С. Система синтеза белорусской речи по тексту / Ю.С. Гецевич, Б.М. Лобанов // Речевые технологии. – 2010. – № 1. – С. 91–100.
8. Гецевіч, Ю.С. Кампаненты для розных платформаў сінтэзатара маўлення па тэксце для інтэлектуальных сістэм / Ю.С. Гецевіч, Д.А. Пакладок, Д.В. Брэк // Открытые семантические технологии проектирования интеллектуальных систем = Open Semantic Technologies for Intelligent Systems (OSTIS–2013) : материалы III Междунар. науч.-техн. конф. (Минск, 21–23 февр. 2013 г.) / редкол. : В.В. Голенков (отв. ред.) [и др.]. – Минск : БГУИР, 2013 г. – С. 375–382.
9. Кароткая граматыка беларускай мовы. У 2 ч. Ч. 1. Фаналогія. Марфалогія. Марфалогія. – Мінск : Беларус. навука, 2007. – 351 с.
10. Беларуская граматыка. У 2 ч. Ч. 1. Фаналогія. Арфаграфія. Марфалогія. Словаўтварэнне. Націск / АН БССР, Інстытут мовазнаўства імя Я.Коласа. – Мінск : Навука і тэхніка, 1985. – С 117–133.
11. Сучасная беларуская мова. Фанетыка. Фаналогія. Арфаэпія. Графіка. Арфаграфія : вучэбна-метадычны комплекс для студэнтаў 1-га курса спецыяльнасцей Г.10.02.01; Г.02.01.00 «Беларуская мова і літаратура». – Мінск : БДУ, 2002. – 144 с.
12. Янкоўскі, Ф.М. Беларускае літаратурнае вымаўленне / Ф.М. Янкоўскі. – Мінск : Народная асвета, 1976. – 91 с.
13. Закон Рэспублікі Беларусь, 23 ліпеня 2008 г., № 420-З. Аб Правілах беларускай арфаграфіі і пунктуацыі [Электронны рэсурс]. – 2013. – Рэжым доступу : <http://www.pravo.by/main.aspx?guid=3871&p0=H10800420&p2>. – Дата доступу : 02.07.2013.
14. Hetsevich, Y. Overview of Belarusian And Russian dictionaries and their adaptation for NooJ / Y. Hetsevich, S. Hetsevich // Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 Intern. Conf. / eds. Vučković Kristina, Bekavac Božo, Silberstein Max. – Newcastle : Cambridge Scholars Publishing, 2012. – P. 29–40.

Паступіла 25.11.2013

*Аб'яднаны інстытут праблем
інфарматыкі НАН Беларусі,
Мінск, Сурганава, 6
e-mail: yury.hetsevich@gmail.com,
lobanov@newman.bas-net.by,
dima.pokladok@gmail.com.*

Yu.S. Hetsevich, B.M. Lobanov, D.A. Pokladok

PHONETIC AND ALLOPHONIC TEXT PROCESSING IN BELARUSIAN AND RUSSIAN SPEECH SYNTHESIZER FOR MOBILE PLATFORMS

The article describes methods of «grapheme – phoneme» and «phoneme – allophone» conversions for Belarusian and Russian speech synthesis. For speech synthesizers on mobile platforms, the rule-based method has been selected. The article describes text processing algorithms at the input phase and rules for «grapheme – phoneme» and «phoneme – allophone» conversions, providing allophone chains for an acoustic processor. The developed algorithms were evaluated with respect to the reference-class stationary Multiphone text-to-speech synthesis system. The test data demonstrates a high accuracy level: over 80% for word processing, over 98% for phoneme processing and over 96% for allophone processing (for Belarusian and Russian languages). The subsequent work will be devoted to the improvement of practical realization of the developed algorithms.

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

УДК 534.26

Г.Ч. Шушкевич, Н.Н. Киселева

ЭКРАНИРОВАНИЕ ЗВУКОВОГО ПОЛЯ ПЛОСКИМ УПРУГИМ СЛОЕМ И ТОНКОЙ НЕЗАМКНУТОЙ СФЕРИЧЕСКОЙ ОБОЛОЧКОЙ

Рассматривается аналитико-численный алгоритм решения граничной задачи, описывающей процесс проникновения звукового поля сферического излучателя, который расположен внутри тонкой незамкнутой сферической оболочки, через плоский упругий слой. Численно исследуется влияние некоторых параметров задачи на значение коэффициента ослабления (экранирования) звукового поля.

Введение

Исследование распространения звуковых волн в упругих средах имеет практическое приложение в медицинской диагностике мягких биологических тканей, гидроакустике, сейсмологии и физической акустике [1–5]. Библиография по решению задач рассеяния весьма обширна. Рассмотрим лишь некоторые работы, относящиеся к данной теме исследования.

В статье [6] методом ступенчатой аппроксимации исследуется влияние количества разбиений (слоев) среды, обладающей упругими свойствами, на дисперсионные и амплитудно-частотные характеристики поверхностных волн. Описание параллельного алгоритма рассеяния звукового поля точечного источника в упругой среде как приложения для задач сейсмологии предлагается в работе [7]. В статье [8] авторы представляют численное исследование дисперсионных характеристик некоторых симметричных и асимметричных моделей для упругих многослойных пакетов с произвольной анизотропией каждого слоя. Влияние физических характеристик упругих слоев среды на рассеяние плоской звуковой волны анализируется методом матриц отражения и прохождения в [9]. В [10] рассматриваются сдвиговые волны в слоистых средах с плоскими границами раздела, дается общее решение для амплитуд сдвиговых волн, приводятся результаты расчета модуля сдвига для структур с меняющимся количеством слоев в зависимости от вида неоднородности. В статье [11] матричным методом находится коэффициент прохождения акустической волны через многослойную пьезоэлектрическую эластичную структуру. В работе [12] методом конечных элементов исследуется рассеяние акустической волны в дальней зоне на упругом объекте в трехслойной среде. Аналогичным методом в статье [13] решается задача дифракции плоской акустической волны через многослойную эластичную структуру, находятся коэффициенты прохождения и отражения. В [14] рассматривается простая акустическая геометрическая система, состоящая из источника звука и тонкой бесконечной упругой пластины, для исследования задачи используется спектральный анализ линейных операторов и находится коэффициент отражения звуковой волны. В [15] описывается аналитическое решение, основанное на матричном методе распространения акустического поля, порождаемого одним источником, в горизонтально-слоистой вязкоупругой среде.

В настоящей работе построено точное осесимметричное решение задачи о проникновении звукового поля через плоский упругий слой. В качестве источника поля рассматривается сферический излучатель, расположенный внутри тонкой незамкнутой сферической оболочки. С помощью соответствующих теорем сложения решение поставленной краевой задачи сведено к решению парных сумматорных уравнений по полиномам Лежандра, которые преобразуются к бесконечной системе линейных алгебраических уравнений второго рода с вполне непрерывным оператором. Исследуется влияние некоторых параметров задачи на значение коэффициента ослабления (экранирования) звукового поля.

1. Постановка задачи и представление решения

Пусть все пространство R^3 разделено плоскостями $S_0(z=h_1)$ и $S_1(z=h_1+h_2)$ на три области $D_0(z < h_1)$, $D_2(h_1 < z < h_1+h_2)$, $D_1(z > h_1+h_2)$ (рис. 1). В области D_0 находится идеально тонкая незамкнутая сферическая оболочка Γ_1 , расположенная на сфере Γ радиуса a с центром в точке O . Область пространства, ограниченную сферой Γ , обозначим через $D_0^{(0)}(0 \leq r < a)$ и $D_0^{(1)} = D_0 \setminus (D_0^{(0)} \cup \Gamma)$. Здесь h_1 – расстояние между точками O и O_1 , h_2 – расстояние между плоскостями S_0 и S_1 .

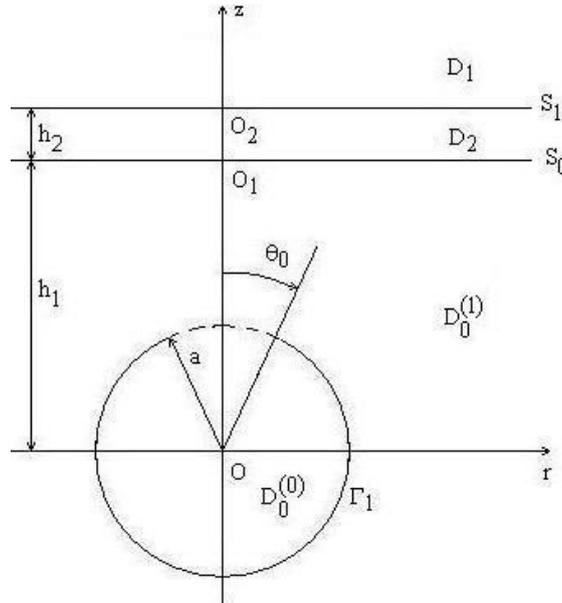


Рис. 1. Геометрия задачи

В точке O расположен точечный излучатель звуковых волн, колеблющихся с круговой частотой ω . Области $D_j, j = \overline{0, 1}$, заполнены материалом, в котором не распространяются сдвиговые волны. Плотность среды и скорость звука в области D_j обозначим соответственно через $\tilde{\rho}_j, c_j, j = \overline{0, 1}$. Область D_2 – плоский упругий слой. Под воздействием звукового поля упругий слой совершает колебания, его деформация определяется вектором смещения \vec{u} , который удовлетворяет уравнению Ламе [16]:

$$\tilde{\mu} \Delta \vec{u} + (\tilde{\lambda} + \tilde{\mu}) \text{grad div } \vec{u} + \omega^2 \tilde{\rho} \vec{u} = 0, \tag{1}$$

где $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ – оператор Лапласа; $\tilde{\lambda}, \tilde{\mu}$ – коэффициенты Ламе; $\tilde{\rho}$ – плотность материала упругой среды.

Для решения задачи свяжем с точкой O сферические координаты $\{r, \theta, \varphi\}$:

$$x = r \cos \varphi \sin \theta, \quad y = r \sin \varphi \sin \theta, \quad z = r \cos \theta,$$

где $0 \leq r < \infty, 0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi$, и цилиндрические координаты $\{\rho, \varphi, z\}$:

$$x = \rho \cos \varphi, \quad y = \rho \sin \varphi, \quad z = z,$$

где $0 \leq \rho < \infty, 0 \leq \varphi \leq 2\pi, -\infty < z < \infty$.

Тонкая незамкнутая сферическая оболочка Γ_1 и плоскости $S_j, j = \overline{0, 1}$, описываются следующим образом:

$$\Gamma_1 = \{r = a, \theta_0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi\}, \quad S_0 = \{z = h_1, 0 \leq \rho \leq \infty, 0 \leq \varphi \leq 2\pi\}, \\ S_1 = \{z = h_1 + h_2, 0 \leq \rho \leq \infty, 0 \leq \varphi \leq 2\pi\}.$$

Обозначим через p_c давление звукового поля точечного излучателя; $p_0^{(0)}$ – давление звукового поля, отраженного от границы Γ_1 в области $D_0^{(0)}$; $p_0 = p_0^{(1)} + p_0^{(2)}$ – суммарное давление рассеянного звукового поля в области $D_0^{(1)}$; $p_0^{(1)}$ – давление звукового поля, отраженного от границы Γ_1 в области $D_0^{(1)}$; $p_0^{(2)}$ – давление звукового поля, отраженного от границы S_0 в области $D_0^{(1)}$; p_1 – давление рассеянного звукового поля в области D_1 .

Реальное смещение и звуковое давление вычисляются по формулам [17]

$$\vec{U} = \text{Re}(\vec{u} e^{-i\omega t}), \quad P_j = \text{Re}(p_j e^{-i\omega t}), \quad j = 0, 1,$$

где i – мнимая единица.

В установившемся режиме колебаний давления рассеянных звуковых полей $p_0^{(j)}, j = \overline{0, 2}$, и p_1 удовлетворяют уравнению Гельмгольца [17, 18]

$$\Delta p_0^{(j)} + k_0^2 p_0^{(j)} = 0 \text{ в } D_0; \quad \Delta p_1 + k_1^2 p_1 = 0 \text{ в } D_1,$$

где $k_0 = \omega / c_0, k_1 = \omega / c_1$ – волновые числа.

В случае распространения малых возмущений в упругом теле для установившегося режима движения частиц тела вектор смещения определяется по формуле (осесимметричная задача) [6]

$$\vec{u} = \text{grad}\psi + \text{rot}\left(-\frac{\partial\Phi}{\partial\rho}\vec{e}_\varphi\right), \quad (2)$$

где функции ψ, Φ удовлетворяют уравнениям Гельмгольца

$$\Delta\psi + k_t^2\psi = 0, \quad k_t = \omega / \tilde{n}_t, \quad \tilde{n}_t = \sqrt{(\tilde{\lambda} + 2\tilde{\mu}) / \tilde{\rho}}; \\ \Delta\Phi + k_t^2\Phi = 0, \quad k_t = \omega / c_t, \quad c_t = \sqrt{\tilde{\mu} / \tilde{\rho}}.$$

Здесь \tilde{n}_t и c_t – скорости распространения продольных и поперечных упругих волн соответственно.

В цилиндрической системе координат компоненты вектора смещения связаны с функциями ψ, Φ соотношениями

$$u_\rho = \frac{\partial\psi}{\partial\rho} + \frac{\partial^2\Phi}{\partial\rho\partial z}, \quad u_z = \frac{\partial\psi}{\partial z} + \frac{\partial^2\Phi}{\partial z^2} + k_t^2\Phi. \quad (3)$$

Решение дифракционной задачи сводится к нахождению вектора смещения \vec{u} и давлений звуковых полей $p_0^{(j)}, j = \overline{0, 2}$, и p_1 , которые удовлетворяют:

– граничному условию на поверхности сферической оболочки Γ_1 (акустически жесткой оболочки):

$$\frac{\partial}{\partial\vec{n}}(p_c + p_0^{(0)})\Big|_{\Gamma_1} = 0, \quad (4)$$

где \vec{n} – нормаль к поверхности Γ_1 ;

– граничным условиям взаимодействия звукового поля с упругим слоем на плоскости S_j , $j = \overline{0, 1}$ [18]:

$$\omega^{-2} \tilde{\rho}_j^{-1} \frac{\partial p_j}{\partial \vec{n}} \Big|_{S_j} = (\vec{u}, \vec{n}) \Big|_{S_j}, \quad T(\vec{u}) \Big|_{S_j} = -p_j \vec{n} \Big|_{S_j},$$

где $T(\vec{u}) = 2\tilde{\mu} \frac{\partial \vec{u}}{\partial z} + \tilde{\lambda} \vec{e}_z \operatorname{div} \vec{u} + \tilde{\mu} [\vec{e}_z, \operatorname{rot} \vec{u}]$, $\vec{n} = \vec{e}_z$ – нормаль к поверхности S_j ,

или

$$u_z \Big|_{S_j} = \omega^{-2} \tilde{\rho}_j^{-1} \frac{\partial p_j}{\partial z} \Big|_{S_j}, \quad \frac{\partial u_\rho}{\partial z} + \frac{\partial u_z}{\partial \rho} \Big|_{S_j} = 0, \quad (2\tilde{\mu} + \tilde{\lambda}) \frac{\partial u_z}{\partial z} + \tilde{\lambda} \left(\frac{u_\rho}{\rho} + \frac{\partial u_\rho}{\partial \rho} \right) \Big|_{S_j} = -p_j \Big|_{S_j}, \quad (5)$$

и условию на бесконечности [19]

$$\lim_{M \rightarrow \infty} r \left(\frac{\partial p_j(M)}{\partial r} - ik_j p_j(M) \right) = 0, \quad j = \overline{0, 1}, \quad (6)$$

где M – произвольная точка пространства.

Потребуем также выполнения условия непрерывности давлений на открытой части сферической оболочки $\Gamma \setminus \Gamma_1$:

$$(p_c + p_0^{(0)}) \Big|_{\Gamma \setminus \Gamma_1} = (p_0^{(1)} + p_0^{(2)}) \Big|_{\Gamma \setminus \Gamma_1}, \quad (7)$$

и нормальной производной на поверхности сферы Γ :

$$\frac{\partial}{\partial \vec{n}} (p_c + p_0^{(0)}) \Big|_{\Gamma} = \frac{\partial}{\partial \vec{n}} (p_0^{(1)} + p_0^{(2)}) \Big|_{\Gamma}, \quad (8)$$

где \vec{n} – нормаль к поверхности Γ .

Давление исходного звукового поля представим в виде выражения [19]

$$p_c(r, \theta) = P \frac{e^{ik_0 r}}{r} = ik_0 P h_0^{(1)}(kr) = P \sum_{n=0}^{\infty} f_n h_n^{(1)}(k_0 r) P_n(\cos \theta), \quad f_n = ik_0 \delta_{0n}, \quad (9)$$

где $h_n^{(1)}(x)$ – сферические функции Ханкеля; $P_n(\cos \theta)$ – полиномы Лежандра [20]; δ_{0k} – символ Кронекера; P – const [17].

Давления рассеянного звукового поля и функции ψ , Φ представим в виде суперпозиции базисных решений уравнения Гельмгольца в сферических и цилиндрических координатах [19, 21], учитывая условие на бесконечности (6):

$$p_0^{(0)}(r, \theta) = P \sum_{n=0}^{\infty} c_n j_n(k_0 r) P_n(\cos \theta) \quad \text{в } D_0^{(0)}; \quad (10)$$

$$p_0^{(1)}(r, \theta) = P \sum_{n=0}^{\infty} x_n h_n^{(1)}(k_0 r) P_n(\cos \theta), \quad p_0^{(2)}(\rho, z) = P \int_0^{\infty} y(\lambda) J_0(\lambda \rho) e^{v_0(z-h_1)} \lambda d\lambda \quad \text{в } D_0^{(1)}; \quad (11)$$

$$p_1(\rho, z) = P \int_0^{\infty} d(\lambda) J_0(\lambda \rho) e^{-v_1(z-h_1-h_2)} \lambda d\lambda \quad \text{в } D_1; \quad (12)$$

$$\psi(\rho, z) = P \int_0^{\infty} \left(a(\lambda) e^{-v_i(z-h_1)} + b(\lambda) e^{v_i(z-h_1-h_2)} \right) J_0(\lambda \rho) \lambda d\lambda; \quad (13)$$

$$\Phi(\rho, z) = P \int_0^{\infty} \left(\tilde{a}(\lambda) e^{-v_i(z-h_1)} + \tilde{b}(\lambda) e^{v_i(z-h_1-h_2)} \right) J_0(\lambda \rho) \lambda d\lambda, \quad (14)$$

где $J_n(x)$ – сферические функции Бесселя первого рода; $J_0(x)$ – функции Бесселя первого рода [19–21]; $v_j = \sqrt{\lambda^2 - k_j^2}$, $-\pi/2 \leq \arg v_j < \pi/2$, $j = 0, 1$; $v_\ell = \sqrt{\lambda^2 - k_\ell^2}$, $-\pi/2 \leq \arg v_\ell < \pi/2$; $v_t = \sqrt{\lambda^2 - k_t^2}$, $-\pi/2 \leq \arg v_t < \pi/2$.

Неизвестные функции $a(\lambda), b(\lambda), \tilde{a}(\lambda), \tilde{b}(\lambda), y(\lambda), d(\lambda)$ и коэффициенты c_n, x_n подлежат определению из граничных условий (4), (5), (7), (8).

2. Выполнение граничных условий

Для выполнения граничных условий (4), (7), (8) представим функцию $p_0^{(2)}(\rho, z)$ через сферические волновые функции, используя формулу [21], связывающую цилиндрические и сферические волновые функции:

$$J_0(\lambda \rho) e^{vz} = \sum_{n=0}^{\infty} (-i)^n (2n+1) P_n \left(\frac{iv}{k} \right) j_n(kr) P_n(\cos \theta), \quad v = \sqrt{\lambda^2 - k^2}, \quad -\pi/2 \leq \arg v < \pi/2.$$

Тогда

$$p_0^{(2)}(r, \theta) = P \sum_{n=0}^{\infty} T_n j_n(k_0 r) P_n(\cos \theta), \quad T_n = (-i)^n (2n+1) \int_0^{\infty} y(\lambda) P_n \left(\frac{iv_0}{k_0} \right) e^{-v_0 h} \lambda d\lambda. \quad (15)$$

Согласно представлениям (9)–(11), (15) граничное условие (8) с учетом условия ортогональности полиномов Лежандра на отрезке $[0; \pi]$ примет вид

$$f_n \frac{d}{d\xi_0} h_n^{(1)}(\xi_0) + c_n \frac{d}{d\xi_0} j_n(\xi_0) = x_n \frac{d}{d\xi_0} h_n^{(1)}(\xi_0) + T_n \frac{d}{d\xi_0} j_n(\xi_0), \quad \xi_0 = k_0 a, \quad n = 0, 1, \dots \quad (16)$$

Выполним граничное условие (4) на поверхности сферической оболочки Γ и условие непрерывности (7). В полученных уравнениях исключим коэффициенты c_n , используя представление (16), и получим парные сумматорные уравнения по полиномам Лежандра вида

$$\begin{cases} \sum_{n=0}^{\infty} \frac{x_n - f_n}{\frac{d}{d\xi_0} j_n(\xi_0)} P_n(\cos \theta) = 0, & 0 \leq \theta < \theta_0; \\ \sum_{n=0}^{\infty} x_n \frac{d}{d\xi_0} h_n^{(1)}(\xi_0) P_n(\cos \theta) = - \sum_{n=0}^{\infty} T_n \frac{d}{d\xi_0} j_n(\xi_0) P_n(\cos \theta), & \theta_0 < \theta \leq \pi. \end{cases} \quad (17)$$

Преобразуем парные сумматорные уравнения (17). Для этого введем в рассмотрение новые коэффициенты X_n по формуле

$$x_n = X_n \frac{d}{d\xi_0} j_n(\xi_0) + f_n, \quad n = 0, 1, \dots, \quad (18)$$

малый параметр g_n по формуле

$$g_n = 1 + \frac{4i\xi_0^3}{2n+1} \frac{d}{d\xi_0} j_n(\xi_0) \frac{d}{d\xi_0} h_n^{(1)}(\xi_0), \quad g_n = O(n^{-2}),$$

и сделаем замену $\theta = \pi - \tilde{\theta}$, $\theta_0 = \pi - \tilde{\theta}_0$, $\tilde{X}_n = (-1)^n X_n$.

В результате парные сумматорные уравнения (17) примут вид

$$\begin{cases} \sum_{n=0}^{\infty} (2n+1)(1-g_n) \tilde{X}_n P_n(\cos \tilde{\theta}) = \sum_{n=0}^{\infty} (-1)^n (2n+1) (\tilde{f}_n + \tilde{T}_n) P_n(\cos \tilde{\theta}), & 0 \leq \tilde{\theta} < \tilde{\theta}_0; \\ \sum_{n=0}^{\infty} \tilde{X}_n P_n(\cos \tilde{\theta}) = 0, & \tilde{\theta}_0 < \tilde{\theta} \leq \pi, \end{cases} \quad (19)$$

где

$$\tilde{T}_n = 4i\xi_0^3 T_n \frac{d}{d\xi_0} j_n(\xi_0) / (2n+1), \quad \tilde{f}_n = 4i\xi_0^3 f_n \frac{d}{d\xi_0} h_n^{(1)}(\xi_0) / (2n+1). \quad (20)$$

Используя интегральные представления для полиномов Лежандра, парные сумматорные уравнения (19) преобразуются к бесконечной системе линейных алгебраических уравнений (СЛАУ) второго порядка с вполне непрерывным оператором [22, 23]:

$$\tilde{X}_n - \sum_{k=0}^{\infty} g_k R_{nk} \tilde{X}_k = \sum_{k=0}^{\infty} (-1)^k (\tilde{T}_k + \tilde{f}_k) R_{nk}, \quad n = 0, 1, \dots, \quad (21)$$

где

$$R_{nk} = \frac{1}{\pi} \left[\frac{\sin(n-k)(\pi-\theta_0)}{n-k} - \frac{\sin(n+k+1)(\pi-\theta_0)}{n+k+1} \right], \quad \left. \frac{\sin(n-k)(\pi-\theta_0)}{n-k} \right|_{n=k} = \pi - \theta_0.$$

Для выполнения граничных условий (5) представим функцию $p_0^{(1)}(r, \theta)$ через цилиндрические волновые функции, используя формулу [21]

$$h_n^{(1)}(kr) P_n(\cos \theta) = \int_0^{\infty} \frac{i^{-n-1}}{kv} P_n\left(\frac{iv}{k}\right) J_0(\lambda \rho) e^{-vz} \lambda d\lambda, \quad v = \sqrt{\lambda^2 - k^2}, \quad -\pi/2 \leq \arg v < \pi/2, \quad z > 0.$$

Тогда

$$p_0^{(1)}(\rho, z) = P \int_0^{\infty} Z(\lambda) J_0(\lambda \rho) e^{-v_0 z} \lambda d\lambda, \quad (22)$$

где

$$Z(\lambda) = \frac{1}{k_0 v_0} \sum_{n=0}^{\infty} i^{-n-1} P_n\left(\frac{iv_0}{k_0}\right) x_n. \quad (23)$$

Принимая во внимание представления (11)–(14), (22) и выполняя граничные условия (5), получим СЛАУ вида

$$M(\lambda) \cdot V(\lambda) = F(\lambda) \cdot Z(\lambda), \quad (24)$$

где

$$M(\lambda) = \begin{pmatrix} m_{11}(\lambda) & m_{12}(\lambda) & m_{13}(\lambda) & m_{14}(\lambda) & 1 & 0 \\ m_{21}(\lambda) & m_{22}(\lambda) & m_{23}(\lambda) & m_{24}(\lambda) & 0 & 0 \\ m_{31}(\lambda) & m_{32}(\lambda) & m_{33}(\lambda) & m_{34}(\lambda) & m_{35}(\lambda) & 0 \\ m_{41}(\lambda) & m_{42}(\lambda) & m_{43}(\lambda) & m_{44}(\lambda) & 0 & 1 \\ m_{51}(\lambda) & m_{52}(\lambda) & m_{53}(\lambda) & m_{54}(\lambda) & 0 & 0 \\ m_{61}(\lambda) & m_{62}(\lambda) & m_{63}(\lambda) & m_{64}(\lambda) & 0 & m_{66}(\lambda) \end{pmatrix}, \quad V(\lambda) = \begin{pmatrix} a(\lambda) \\ b(\lambda) \\ \tilde{a}(\lambda) \\ \tilde{b}(\lambda) \\ y(\lambda) \\ d(\lambda) \end{pmatrix}, \quad F(\lambda) = \begin{pmatrix} f_1(\lambda) \\ 0 \\ f_3(\lambda) \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

$$\begin{aligned}
m_{11}(\lambda) &= (2\tilde{\mu} + \tilde{\lambda})v_\ell^2 - \tilde{\lambda}\lambda^2, \quad m_{12}(\lambda) = \left[(2\tilde{\mu} + \tilde{\lambda})v_\ell^2 - \tilde{\lambda}\lambda^2 \right] e^{-v_\ell h_2}, \\
m_{13}(\lambda) &= (2\tilde{\mu} + \tilde{\lambda})(-v_t^3 - v_t k_t^2) + \tilde{\lambda}\lambda^2 v_t, \quad m_{14}(\lambda) = \left[(2\tilde{\mu} + \tilde{\lambda})(v_t^3 + v_t k_t^2) - \tilde{\lambda}\lambda^2 v_t \right] e^{-v_t h_2}, \\
m_{21}(\lambda) &= -2v_\ell, \quad m_{22}(\lambda) = 2v_\ell e^{-v_\ell h_2}, \quad m_{23}(\lambda) = 2v_t^2 + k_t^2, \quad m_{24}(\lambda) = \left[2v_t^2 + k_t^2 \right] e^{-v_t h_2}, \\
m_{31}(\lambda) &= -v_\ell, \quad m_{32}(\lambda) = v_\ell e^{-v_\ell h_2}, \quad m_{33}(\lambda) = v_t^2 + k_t^2, \quad m_{34}(\lambda) = \left[v_t^2 + k_t^2 \right] e^{-v_t h_2}, \\
m_{35}(\lambda) &= -\omega^{-2} \tilde{\rho}_0^{-1} v_0, \quad m_{41}(\lambda) = \left[(2\tilde{\mu} + \tilde{\lambda})v_\ell^2 - \tilde{\lambda}\lambda^2 \right] e^{-v_\ell h_2}, \quad m_{42}(\lambda) = (2\tilde{\mu} + \tilde{\lambda})v_\ell^2 - \tilde{\lambda}\lambda^2, \\
m_{43}(\lambda) &= \left[(2\tilde{\mu} + \tilde{\lambda})(-v_t^3 - v_t k_t^2) + \tilde{\lambda}\lambda^2 v_t \right] e^{-v_t h_2}, \quad m_{44}(\lambda) = (2\tilde{\mu} + \tilde{\lambda})(v_t^3 + v_t k_t^2) - \tilde{\lambda}\lambda^2 v_t, \\
m_{51}(\lambda) &= -2v_\ell e^{-v_\ell h_2}, \quad m_{52}(\lambda) = 2v_\ell, \quad m_{53}(\lambda) = \left[2v_t^2 + k_t^2 \right] e^{-v_t h_2}, \quad m_{54}(\lambda) = 2v_t^2 + k_t^2, \\
m_{61}(\lambda) &= -v_\ell e^{-v_\ell h_2}, \quad m_{62}(\lambda) = v_\ell, \quad m_{63}(\lambda) = \left[v_t^2 + k_t^2 \right] e^{-v_t h_2}, \quad m_{64}(\lambda) = v_t^2 + k_t^2, \\
m_{66}(\lambda) &= \omega^{-2} \tilde{\rho}_1^{-1} v_1, \quad f_1(\lambda) = -e^{-v_0 h_1}, \quad f_3(\lambda) = -\omega^{-2} \tilde{\rho}_1^{-1} v_1 e^{-v_0 h_1}.
\end{aligned}$$

Решая систему (24), найдем представление для функции $y(\lambda)$:

$$y(\lambda) = |M_5(\lambda)| Z(\lambda) / |M(\lambda)|, \quad (25)$$

где $|M(\lambda)|$ – определитель матрицы $M(\lambda)$; $|M_5(\lambda)|$ – определитель матрицы $M_5(\lambda)$; $M_5(\lambda)$ – матрица $M(\lambda)$, в которой пятый столбец заменен на вектор $F(\lambda)$.

Подставим функцию $y(\lambda)$ из (25) в (15), учитывая представление (20), и установим связь между коэффициентами \tilde{T}_k и функцией $Z(\lambda)$. В полученном выражении заменим функцию $Z(\lambda)$ представлением (23), коэффициенты x_n представлением (18) и получим связь между коэффициентами \tilde{T}_k и \tilde{X}_p :

$$\tilde{T}_k = \sum_{p=0}^{\infty} S_{pk} \tilde{X}_p + \tilde{f}_k, \quad k=0, 1, 2, \dots, \quad (26)$$

где

$$S_{pk} = 4\xi_0^3 (-1)^k i^{k+p} \frac{d}{d\xi_0} j_p(\xi_0) \frac{d}{d\xi_0} j_k(\xi_0) \int_0^{\infty} \frac{|M_5(\lambda)|}{k_0 v_0 |M(\lambda)|} P_p \left(\frac{iv_0}{k_0} \right) P_k \left(\frac{iv_0}{k_0} \right) e^{-v_0 h_1} \lambda d\lambda;$$

$$\tilde{f}_k = 4\xi_0^3 i(-i)^{k+1} \frac{d}{d\xi_0} j_k(\xi_0) \int_0^{\infty} \frac{|M_5(\lambda)|}{v_0 |M(\lambda)|} P_k \left(\frac{iv_0}{k_0} \right) e^{-v_0 h_1} \lambda d\lambda.$$

Преобразуем правую часть системы (21). Для этого исключим из правой части коэффициенты \tilde{T}_k с помощью представления (26) и получим бесконечную СЛАУ второго порядка относительно \tilde{X}_n :

$$\tilde{X}_n - \sum_{k=0}^{\infty} (g_k R_{nk} - \alpha_{nk}) \tilde{X}_k = \sum_{k=0}^{\infty} \left(\tilde{f}_k + (-1)^k \tilde{f}_k \right) R_{nk}, \quad n=0, 1, 2, \dots, \quad (27)$$

$$\alpha_{nk} = \sum_{p=0}^{\infty} (-1)^p R_{np} S_{kp}.$$

Находим связь между функцией $d(\lambda)$, входящей в представление вторичного давления $p_1(\rho, z)$ в области D_1 , и коэффициентами \tilde{X}_n – решениями системы (27). Из системы (24) следует, что

$$d(\lambda) = |M_6(\lambda)| Z(\lambda) / |M(\lambda)|, \quad (28)$$

где $|M_6(\lambda)|$ – определитель матрицы $M_6(\lambda)$; $M_6(\lambda)$ – матрица $M(\lambda)$, в которой шестой столбец заменен на вектор $F(\lambda)$.

Учитывая связь функции $Z(\lambda)$ с коэффициентами \tilde{X}_n (см. формулы (18), (23)), имеем

$$d(\lambda) = \frac{|M_6(\lambda)|}{|M(\lambda)| k_0 v_0} \sum_{p=0}^{\infty} i^{-p-1} P_p \left(\frac{iv_0}{k_0} \right) \left((-1)^p \tilde{X}_p \frac{d}{d\xi_0} j_p(\xi_0) + f_p \right).$$

Коэффициент ослабления (экранирования) звукового поля в области D_1 вычисляем по формуле

$$K(\rho, z) = |p_1(\rho, z)| / |p_c|, \quad z > h_1 + h_2.$$

Если незамкнутая сферическая оболочка Γ_1 отсутствует, то решение задачи сводится к решению системы (24), а величина $Z(\lambda) = 1/v_0$. В этом случае функция $d(\lambda)$, входящая в представление вторичного давления $p_1(\rho, z)$ в области D_1 , вычисляется по формуле

$$d(\lambda) = |M_6(\lambda)| / |M(\lambda)|, \quad (29)$$

где $|M_6(\lambda)|$ – определитель матрицы $M_6(\lambda)$; $M_6(\lambda)$ – матрица $M(\lambda)$, в которой шестой столбец заменен на вектор $F(\lambda)/v_0$.

3. Вычислительный эксперимент

Используя систему компьютерной математики Mathcad [24], были проведены вычисления коэффициента ослабления $K(\rho, z)$ звукового поля в области D_1 для некоторых параметров задачи.

Специальные функции $j_n(x)$, $h_n^{(1)}(x) = j_n(x) + iy_n(x)$ ($y_n(x)$ – сферическая функция Бесселя второго рода), $J_0(x)$ и полиномы Лежандра $P_n(\cos\theta)$ вычислялись с помощью встроенных функций [24]. Производные сферических функций вычислялись с помощью формулы [20, с. 258]

$$\frac{d}{dx} f_n(x) = n f_n(x) / x - f_{n+1}(x), \quad n = 0, 1, 2, \dots$$

Величины $v_j = \sqrt{\lambda^2 - k_j^2}$, $j=0, 1$, $v_\ell = \sqrt{\lambda^2 - k_\ell^2}$, $v_\tau = \sqrt{\lambda^2 - k_\tau^2}$, входящие в представление (11)–(14), вычислялись по формуле

$$v_\tau = \begin{cases} \sqrt{\lambda^2 - k_\tau^2}, & \lambda \geq k_\tau; \\ -i\sqrt{k_\tau^2 - \lambda}, & 0 \leq \lambda < k_\tau. \end{cases}$$

Бесконечная система (27) решалась методом усечения [19]. Для получения достоверного решения конечной системы линейных алгебраических уравнений необходимо проверить обу-

словленность системы. Матрица, соответствующая системе, считается хорошо обусловленной, если число обусловленности матрицы меньше 300 [25, с.150]. Для вычисления числа обусловленности матрицы использовались встроенные функции `cond1` (число обусловленности в норме L_1 [26]), `cond2` (в норме L_2) и `conde` (в евклидовой норме). Вычислительный эксперимент показал, что порядок усечения системы (27) можно взять равным 17 для рассмотренных параметров задачи. Это обеспечивает решение конечной системы (27) с точностью 10^{-4} и число обусловленности не будет превосходить 35. Несобственные интегралы вычислялись приемами, предложенными в [26].

На рис. 2 изображены графики коэффициента ослабления звукового поля $K(0, z)$, $z > h_1 + h_2$, для некоторых значений угла θ_0 при отсутствии незамкнутой сферической оболочки Γ_1 , если область D_0 заполнена воздухом ($\tilde{\rho}_0 = 1,29 \text{ кг/м}^3$, $c_0 = 343 \text{ м/с}$), область D_1 – водой ($\tilde{\rho}_1 = 1000 \text{ кг/м}^3$, $c_1 = 1500 \text{ м/с}$), область D_2 – каучуком ($\tilde{\rho} = 910 \text{ кг/м}^3$, $E = 0,79 \cdot 10^7 \text{ Н/м}^2$ – модуль Юнга, $\nu = 0,46$ – коэффициент Пуассона), $h_1 = 4 \text{ м}$, $h_2 = 0,01 \text{ м}$, $a = 0,2 \text{ м}$, $f = 50 \text{ Гц}$, $\omega = 2\pi f$. Коэффициенты Ламе связаны с модулем Юнга и коэффициентом Пуассона соотношениями

$$\tilde{\lambda} = \nu E / ((1 + \nu)(1 - 2\nu)), \quad \tilde{\mu} = E / (2 + 2\nu).$$

Графики показывают, что при увеличении значения угла θ_0 график коэффициента $K(0, z)$ совпадает с графиком коэффициента $K(0, z)$ в случае отсутствия незамкнутой сферической оболочки Γ_1 . Таким образом, малый угол раствора $\pi - \theta_0$ тонкой сферической оболочки Γ_1 практически не оказывает влияние на коэффициент экранирования.

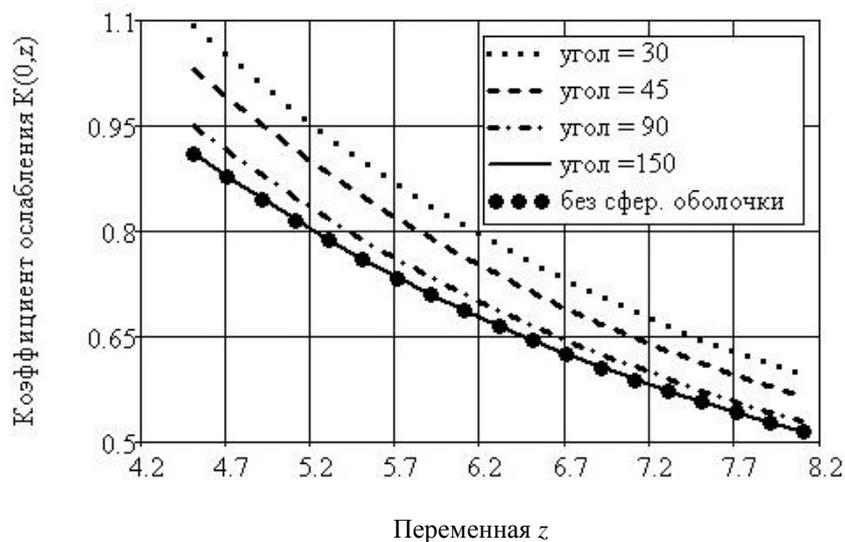


Рис. 2. Графики коэффициента ослабления звукового поля $K(0, z)$ для некоторых значений угла θ_0

На рис. 3 изображены графики коэффициента ослабления звукового поля $K(0, z)$, $z > h_1 + h_2$, если область D_0 заполнена машинным маслом ($\tilde{\rho}_0 = 900 \text{ кг/м}^3$, $c_0 = 1390 \text{ м/с}$), область D_1 – морской водой ($\tilde{\rho}_1 = 1030 \text{ кг/м}^3$, $c_1 = 1510 \text{ м/с}$), глицерином ($\tilde{\rho}_1 = 1260 \text{ кг/м}^3$, $c_1 = 1904 \text{ м/с}$), керосином ($\tilde{\rho}_1 = 800 \text{ кг/м}^3$, $c_1 = 2330 \text{ м/с}$), кирпичом ($\tilde{\rho}_1 = 1500 \text{ кг/м}^3$, $c_1 = 3600 \text{ м/с}$), область D_2 – алюминием ($\tilde{\rho} = 2600 \text{ кг/м}^3$, $E = 65 \cdot 10^9 \text{ Н/м}^2$ – модуль Юнга, $\nu = 0,32$ – коэффициент Пуассона), $h_1 = 2 \text{ м}$, $h_2 = 0,01 \text{ м}$, $a = 0,1 \text{ м}$, $f = 150 \text{ Гц}$, $\theta_0 = \pi / 2$.

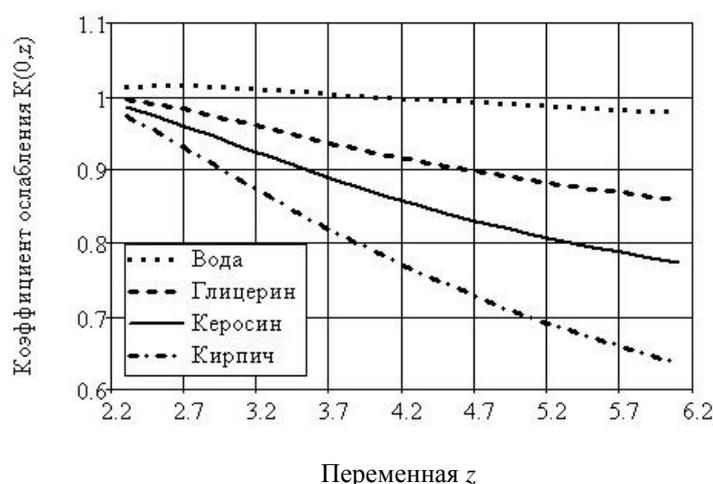


Рис. 3. Графики коэффициента ослабления звукового поля $K(0,z)$ в случае заполнения области D_1 различными веществами

На рис. 4 изображены графики коэффициента ослабления звукового поля $K(0,z)$, $z > h_1 + h_2$, если область D_0 заполнена машинным маслом ($\tilde{\rho}_0 = 900 \text{ кг/м}^3$, $c_0 = 1390 \text{ м/с}$), область D_1 – водой ($\tilde{\rho}_1 = 1000 \text{ кг/м}^3$, $c_1 = 1500 \text{ м/с}$), область D_2 – каучуком ($\tilde{\rho} = 910 \text{ кг/м}^3$, $E = 0,79 \cdot 10^7 \text{ Н/м}^2$, $\nu = 0,46$), стеклом ($\tilde{\rho} = 2500 \text{ кг/м}^3$, $E = 49 \cdot 10^9 \text{ Н/м}^2$, $\nu = 0,2$), вольфрамом ($\tilde{\rho} = 19250 \text{ кг/м}^3$, $E = 39 \cdot 10^{10} \text{ Н/м}^2$, $\nu = 0,29$), иридием ($\tilde{\rho} = 22650 \text{ кг/м}^3$, $E = 52,8 \cdot 10^{10} \text{ Н/м}^2$, $\nu = 0,26$), $h_1 = 3 \text{ м}$, $h_2 = 0,01 \text{ м}$, $a = 0,25 \text{ м}$, $\theta_0 = \pi/2$, $f = 50 \text{ Гц}$.

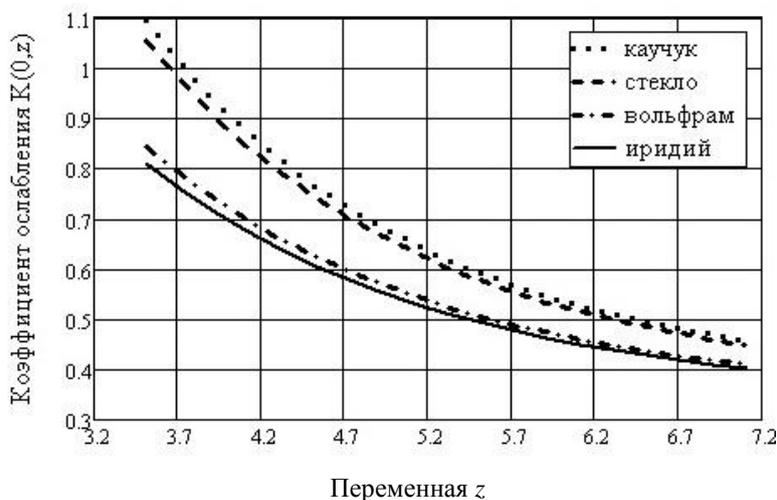


Рис. 4. Графики коэффициента ослабления звукового поля $K(0,z)$ в случае заполнения области D_2 различными веществами

На рис. 5 изображены графики коэффициента ослабления звукового поля $K(0,z)$, $z > h_1 + h_2$, для различных значений частоты исходного звукового поля в случае, если область D_0 заполнена воздухом ($\tilde{\rho}_0 = 1,29 \text{ кг/м}^3$, $c_0 = 343 \text{ м/с}$), область D_1 – азотом ($\tilde{\rho}_1 = 830 \text{ кг/м}^3$, $c_1 = 962 \text{ м/с}$), область D_2 – алюминием ($\tilde{\rho} = 2600 \text{ кг/м}^3$, $E = 65 \cdot 10^9 \text{ Н/м}^2$, $\nu = 0,32$), $h_1 = 4 \text{ м}$, $h_2 = 0,02 \text{ м}$, $a = 0,3 \text{ м}$, $\theta_0 = \pi/2$.

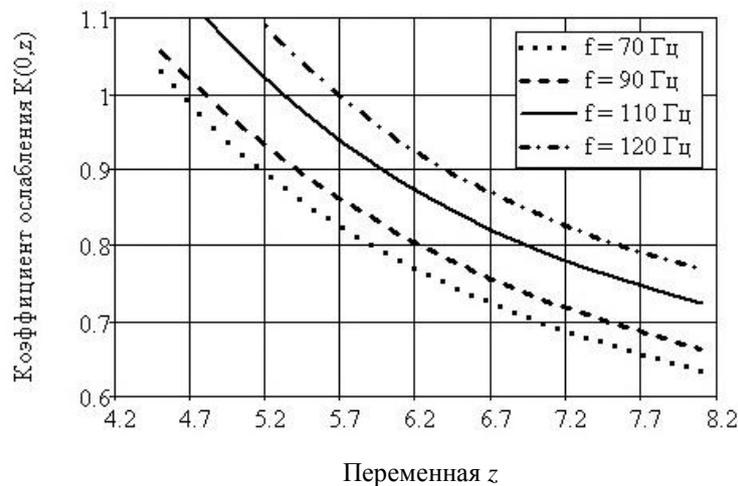


Рис. 5. Графики коэффициента ослабления звукового поля $K(0, z)$ для различных значений частоты исходного поля

На рис. 6 изображены графики коэффициента ослабления звукового поля $K(0, z)$, $z > h_1 + h_2$, в случае отсутствия незамкнутой сферической оболочки Γ_1 для различных значений h_2 , если область D_0 заполнена машинным маслом ($\tilde{\rho}_0 = 900 \text{ кг/м}^3$, $c_0 = 1390 \text{ м/с}$), область D_1 – керосином ($\tilde{\rho}_1 = 800 \text{ кг/м}^3$, $c_1 = 2330 \text{ м/с}$), область D_2 – каучуком ($\tilde{\rho} = 910 \text{ кг/м}^3$, $E = 0,79 \cdot 10^7 \text{ Н/м}^2$, $\nu = 0,46$), $h_1 = 3,5 \text{ м}$, $f = 50 \text{ Гц}$.

Графики показывают, что при уменьшении толщины h_2 упругого слоя практически отсутствует экранирующий эффект.

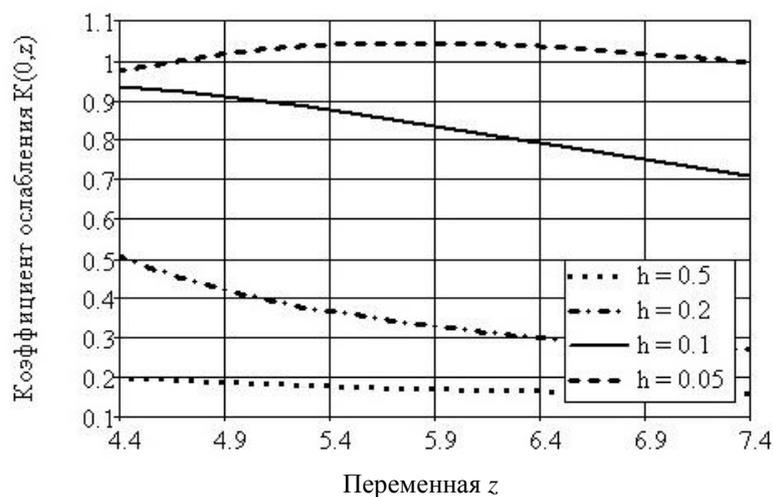


Рис. 6. Графики коэффициента ослабления звукового поля $K(0, z)$ для различных значений толщины упругого слоя в случае отсутствия тонкой сферической оболочки

Заключение

С помощью теоремы сложения для цилиндрических и сферических волновых функций решение поставленной задачи о проникновении звукового поля через плоский упругий слой сведено к решению парных сумматорных уравнений по полиномам Лежандра. Парные уравнения преобразованы к бесконечной системе линейных алгебраических уравнений второго

рода с вполне непрерывным оператором. В качестве источника звукового поля рассматривается сферический излучатель, расположенный внутри тонкой незамкнутой сферической оболочки.

Численно исследовано влияние геометрических параметров задачи, плотности сред, модуля Юнга, коэффициента Пуассона и скорости звука в средах на значение коэффициента ослабления звукового поля. Вычислительные эксперименты показали, что если упругий слой экрана заполнен веществом с большим значением модуля Юнга, эффективность экранирования улучшается. Коэффициент экранирования уменьшается, если скорости распространения звука в средах, в которых не распространяются сдвиговые волны, различные. Коэффициент ослабления поля зависит от круговой частоты звукового источника ω и угла θ_0 : с увеличением частоты ω коэффициент экранирования увеличивается, с увеличением угла θ_0 – уменьшается. Разработанная методика и программное обеспечение могут найти практическое использование при изготовлении звуковых экранов.

Список литературы

1. Бреховских, Л.М. Волны в слоистых средах / Л.М. Бреховских. – М.: Изд-во АН СССР, 1957. – 502 с.
2. Shear wave elasticity imaging: a new ultrasonic technology of medical diagnostic / A.P. Sarvazyan [et al.] // *Ultrasonics Med. Biol.* – 1998. – Vol. 24. – P. 1419–1435.
3. Иванов, Н.И. Инженерная акустика. Теория и практика борьбы с шумом / Н.И. Иванов. – М.: Логос, 2008. – 424 с.
4. Абракитов, В.Э. Многослойная ограждающая панель / В.Э. Абракитов, В.А. Руссова [Электронный ресурс]. – Режим доступа : <http://abraitov.narod.ru/Monograph/6-1.htm>. – Дата доступа : 07.09.2013.
5. IV Всерос. науч.-практ. конф. с международным участием «Защита населения от повышенного шумового воздействия» // [Электронный ресурс]. – Режим доступа : <http://onlinereg.ru/noise2013>. – Дата доступа : 07.10.2013.
6. Поверхностные волны в материалах с функционально-градиентными покрытиями / Е.В. Глушков [и др.] // *Акустический журнал*. – 2012. – № 3. – С. 370–385.
7. Fatyanov, A.G. High-performance modeling acoustic and elastic waves using the parallel Dichotomy Algorithm / A.G. Fatyanov, A.V. Terekhov // *Journal of Computational Physics*. – 2011. – Vol. 230. – P. 699–715.
8. Investigation of Lamb elastic waves in anisotropic multilayered composites applying the Green's matrix / A. Karmazin [et al.] // *Ultrasonics*. – 2011. – Vol. 51. – P. 17–28.
9. An analytical solution for wave-induced seabed response in a multi-layered poro-elastic seabed / X-L. Zhou [et al.] // *Ocean Engineering*. – 2011. – Vol. 38. – P. 119–129.
10. Андреев, В.Р. Стоячие сдвиговые волны в слоистых резиноподобных средах / В.Р. Андреев, Т.Б. Крит, О.А. Сапожников // *Акустический журнал*. – 2010. – Т. 56. – № 5. – С. 579–586.
11. Acoustic wave transmission through piezoelectric structured materials / M. Lam [et al.] // *Ultrasonics*. – 2009. – Vol. 49. – P. 424–431.
12. Ito, K. A domain decomposition solver for acoustic scattering by elastic objects in layered media / K. Ito, Z. Qiao, J. Toivanen // *Journal of Computational Physics*. – 2008. – Vol. 227. – P. 8685–8698.
13. Sadler, J. A ray technique to calculate reflected and transmitted waves in layered media / J. Sadler, R. Gr. Maev // *ScienceDirect*. – 2008. – Vol. 48. – P. 687–696.
14. Pokrovski, A.A. Lax-Phillips theory for sound waves scattering by thin infinite elastic planes / A.A. Pokrovski // *Computers Math. Applic.* – 1997. – Vol. 34. – P. 143–171.
15. Schmidt, H. Efficient numerical solution technique for wave propagation in horizontally stratified environments / H. Schmidt, F.B. Jensen // *Comp. and Maths. with Appls.* – 1985. – Vol. 11. – P. 699–715.
16. Новацкий, В. Теория упругости / В. Новацкий. – М.: Мир, 1970. – 873 с.
17. Шендарев, Е.Л. Излучение и рассеяние звука / Е.Л. Шендарев. – Л.: Судостроение, 1989. – 304 с.

18. Ерофеенко, В.Т. Моделирование двухсторонних граничных условий для акустических волн на упругом экране / В.Т. Ерофеенко // Весці НАН Беларусі. Сер. фіз.-мат. навук. – 2010. – № 4. – С. 76–84.
19. Иванов, Е.А. Дифракция электромагнитных волн на двух телах / Е.А. Иванов. – Минск : Наука и техника, 1968. – 584 с.
20. Справочник по специальным функциям с формулами, графиками и таблицами / под ред. М. Абрамовица, И. Стиган. – М. : Наука, 1979. – 830 с.
21. Ерофеенко, В.Т. Теоремы сложения / В.Т. Ерофеенко. – Минск : Наука и техника, 1989. – 240 с.
22. Шушкевич, Г.Ч. Расчет электростатических полей методом парных, тройных уравнений с использованием теорем сложения / Г.Ч. Шушкевич. – Гродно : ГрГУ, 1999. – 238 с.
23. Резуненко, В.А. Дифракция плоской звуковой волны на сфере с круговым отверстием / В.А. Резуненко // Вісник Харків. нац. універ. ім. В.Н. Каразіна. Сер. «Мат., прик. мат. і мех.». – 2009. – № 850. – С. 71–77.
24. Шушкевич, Г.Ч. Компьютерные технологии в математике. Система Mathcad 14. Ч. 1. / Г.Ч. Шушкевич, С.В. Шушкевич. – Минск : Изд-во Гревцова, 2010. – 287 с.
25. Петров, Ю. Обеспечение достоверности и надежности компьютерных расчетов / Ю. Петров. – СПб. : БХВ – Петербург, 2012. – 160 с.
26. Вержбицкий, В.М. Основы численных методов / В.М. Вержбицкий. – М. : Высшая школа, 2002. – 848 с.

Поступила 11.11. 2013

*Гродненский государственный
университет им. Янки Купалы,
Гродно, Ожешко, 22
e-mail: g_shu@tut.by*

G.Ch. Shushkevich, N.N. Kiselyova

SOUND FIELD SHIELDING BY FLAT ELASTIC LAYER AND THIN UNCLOSED SPHERICAL SHELL

An analytical solution of a boundary problem describing the process of penetration of a sound field of a spherical radiator located inside a thin unclosed spherical shell through a flat elastic layer is constructed. An influence of some parameters of the problem on the value of the attenuation coefficient (screening) of the sound field was studied by using a numerical simulation.

УДК 539.3

О.Л. Швед

ОПРЕДЕЛЕНИЕ ДЕВИАТОРНОГО СЕЧЕНИЯ ПОВЕРХНОСТИ ТЕКУЧЕСТИ ПРИ МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ УПРУГОПЛАСТИЧЕСКОГО ПОВЕДЕНИЯ МАТЕРИАЛОВ

Рассматривается перспективный подход к созданию нелинейных моделей упругопластичности на основе обобщения нелинейных моделей упругости. Для обобщенного материала Мурнагана предлагается способ нахождения вектора нормали к поверхности девиаторного сечения в условиях анизотропии.

Введение

Известно, что состояние теории нелинейно упругопластических материалов в отличие от нелинейно упругих материалов является неудовлетворительным [1]. Об этом свидетельствуют постоянно появляющиеся публикации, направленные на пересмотр основ теории. В упругости такого не наблюдается. Теоретические модели пытаются вывести из экспериментальных фактов, при этом отсутствует общая теоретическая концепция.

В настоящем исследовании взамен концепции гладкого дифференцированного многообразия, которая в упругопластичности работает «за пределами своих возможностей», предлагается использовать либо полностью, либо частично принцип сохранения потенциальной природы упругой деформации.

1. Подход к построению модели упругопластического материала

В работах [2, 3] предложены определяющие соотношения упругопластичности, уточненные по результатам проведенных исследований. Геометрически нелинейная модель упругости [4, 5] обобщается на геометрически нелинейную модель упругопластичности. Транспонированный градиент деформации $\overset{0}{\nabla} \mathbf{R}^T$ в упругом законе заменяется неособенным тензором \mathbf{F}_e , который определяется не в рамках кинематики. Напряжения в упругопластичности для анизотропного материала не связаны с мерой деформации, порождаемой чисто геометрическими построениями, как это получается в упругости [1].

Сохраняется потенциальная природа упругой деформации: требуется существование потенциалов напряжений и их скоростей. Второе требование является обязательным, а первое оказывается желательным, поскольку позволяет описать важное явление роста анизотропии материала в результате пластической деформации. Кстати, само понятие пластической деформации в предлагаемом подходе не определяется в общем случае в силу неоднозначности его определения [6]. Потенциальность в напряжениях ввел Грин (1839) для упругих материалов и для неупругих не использовал. Потенциальность в скоростях напряжений предлагали Хилл (1950), Ключников (1960) и другие исследователи для упругопластических материалов. Требование двойной потенциальности ранее не предлагалась, поскольку в стандартных подходах игнорировался девиатор упругих напряжений. Это требование дает теоретическое определение девиаторных сечений поверхности текучести и, следовательно, ее самой с учетом экспериментальных данных. Экспериментальными методами определить поверхность текучести в пространстве напряжений, как недавно выяснилось, невозможно [1]. При этом в предлагаемом подходе любой гиперупругий закон нелинейной упругости позволяет определить девиаторные сечения поверхности текучести, все кривые пластичности практически совпадают и близки к кривой пластичности А.Ю. Ишлинского.

В настоящей работе используется концепция поверхности текучести в пространстве напряжений. Активный упругопластический процесс представляем попеременным чередованием пластического и упругого состояний элементов деформируемого твердого тела на основа-

нии принципа В.Д. Ключникова [7]. Он состоит в замене данного пути нагружения некоторым, как угодно близким, на участках которого свойства пластичности либо заранее известны, либо постулируются более естественно, чем на исходном пути. Здесь выбирается специальный путь нагружения. Использование указанного принципа наряду с упрощением решения проблемы формулировки определяющих соотношений позволяет смягчить жесткость условия текучести (ввиду несжимаемости материала) с целью улучшения адекватности моделирования проблемных течений при обработке металлов давлением. Упругих и пластических зон, вообще говоря, не будет.

Согласно [1] при классическом подходе существующие теории пластичности, основанные на условии текучести, не описывают опыты Треска по экструзии свинца [8]. В работе [9] для проверки теории при численном моделировании процесса прямого выдавливания форма контейнера выбрана с углом при основании более 90° , в связи с этим проблемное течение в случае прямого угла исключалось. Возможности описания проблемных течений с использованием системы LS-DYNA для геометрически линейного материала Elastic Plastic и геометрически нелинейного Vampan проверялись в работах [10, 11]. Так, при моделировании опыта Треска с практически нулевым трением [10] наблюдались значительные отклонения от эксперимента и появление застойной зоны, которая в опыте, где трение было значительно больше, отсутствовала, а моделирование явления двойного бочкообразования при осадке высоких заготовок [11] показало качественное расхождение с опытными данными. При этом численные данные указанного явления в литературе также отсутствуют. Отметим еще одно проблемное течение – явление «запирания» области высокого давления (переход материалов из пластического состояния в упругое) [12]. Оно состоит в следующем. При сжатии тонкого диска на наковальне Бриджмена [13] начиная с некоторой величины усилия в центре диска образуется и расширяется упругая область и никаким сколь угодно большим давлением не удастся изменить остаточную толщину. Аналогичное явление наблюдается при деформировании материалов в аппарате высокого давления типа наковальни с лункой. С некоторой величины усилия вытекание материала из полости лунки прекращается, т. е. материал в лунке деформируется упруго. Эти явления привычны для практиков, но необычны с точки зрения механики материалов. Возможное их объяснение – невыпуклость поверхности текучести [12]. В силу того что поверхность девиаторного сечения получилась кусочно-вогнутой и подвижность материала была увеличена за счет чередования пластических и упругих состояний элементов деформируемого твердого тела, должны описываться явление «запирания» области высокого давления и опыты Треска по экструзии свинца.

Используется векторное представление симметричных девиаторов второго ранга, скалярное произведение таких «векторов» понимается как двойное скалярное произведение тензоров. Пусть $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ – неподвижный ортонормированный триэдр. Задается ортонормированный тензорный базис пространства девиаторов, два первых вектора из которого имеют вид $\mathbf{W}_1 = (\sqrt{6})^{-1}(\mathbf{E} - 3\mathbf{c}_3\mathbf{c}_3)$, $\mathbf{W}_2 = (\sqrt{2})^{-1}(\mathbf{c}_2\mathbf{c}_2 - \mathbf{c}_1\mathbf{c}_1)$. Девиатор \mathbf{N} в плоскости $\{\mathbf{W}_1, \mathbf{W}_2\}$ представляется в векторном виде $\mathbf{N} = \sum_i w_i \mathbf{W}_i$.

В пластическом состоянии (при течении) материал предполагаем несжимаемым. Обозначим $\nabla \mathbf{v}$ и $\nabla \mathbf{v}^T$ – градиент и транспонированный градиент скорости соответственно. При течении по соотношениям нелинейной упругости $(\overset{0}{\nabla} \mathbf{R}^T)^{\cdot} = \nabla \mathbf{v}^T \cdot \overset{0}{\nabla} \mathbf{R}^T$ при условии несжимаемости $\nabla \cdot \mathbf{v} = 0$ вычисляется тензор скорости упругих напряжений с использованием объективной О-производной: $\overset{\Omega}{\mathbf{T}} = \overset{\Omega}{\dot{\mathbf{T}}} - \mathbf{\Omega} \cdot \mathbf{T} + \mathbf{T} \cdot \mathbf{\Omega}$, где $\overset{\Omega}{\dot{\mathbf{T}}}$ – материальная производная тензора напряжений Коши \mathbf{T} ; $\mathbf{\Omega}$ – косимметричный тензор упругого спина [14].

Поверхность текучести накладывает ограничение на девиаторную составляющую $\mathbf{Q} = \mathbf{Q}(\mathbf{D})$, где $\mathbf{D} = 2^{-1}(\nabla \mathbf{v} + \nabla \mathbf{v}^T)$ – тензор скорости деформаций. Поэтому при задании тензора скорости напряжений шаровая часть удаляется, а вектор \mathbf{Q} проецируется ортогонально

нормированному вектору внешней нормали \mathbf{N} ($\mathbf{N} \cdot \mathbf{N} = 1$) к поверхности девиаторного сечения поверхности текучести, причем берется только независимая от девиатора \mathbf{D} часть его проекции K . Формально это записывается как

$$\overset{\Omega}{\mathbf{T}} = K(\mathbf{Q} - \mathbf{Q} \cdot \mathbf{N}\mathbf{N}), \quad (1)$$

где K – малый скаляр, определяемый в эксперименте по двухосным растяжению и сжатию. Из (1) следует, что в силу потенциальности тензоров $\overset{\Omega}{\mathbf{T}}$ и \mathbf{Q} (первого тензора по модельному предположению, а второго – по построению) девиатор $\mathbf{Q} \cdot \mathbf{N}\mathbf{N}$ должен иметь потенциал. В уравнении (1) и следствии из него заключается, вероятно, суть предлагаемого подхода. Подчеркнем, что девиатор упругих напряжений $\text{dev } \mathbf{T}$, который не зависит от скоростей, здесь не игнорируется (как в классическом подходе), а используется.

При разработке модели материала наряду с техническими трудностями, вызванными громоздкостью используемых соотношений, возникает вопрос об определении вектора нормали к поверхности девиаторного сечения в условиях нетривиальной анизотропии, например, для моноклинного материала. Известно [2, 3], что вектор нормали является одним из собственных векторов оператора $\mathbf{Q} = \mathbf{Q}(\mathbf{D})$. Для вектора нормали предлагается следующий способ его выбора из пяти вычисленных собственных векторов оператора. Из условия потенциальности девиатора $\mathbf{Q} \cdot \mathbf{N}\mathbf{N}$ можно найти в явном виде этот вектор \mathbf{N}_0 для изотропного материала и затем сравнивать его с вычисленными векторами. Ближайший вектор к вектору \mathbf{N}_0 будет искомым. Перейдем к отысканию вектора $\mathbf{N} = \mathbf{N}_0$.

2. Изотропный упругий закон Мурнагана и вспомогательные соотношения

Рассмотрим изотропный упругий материал. Потенциал напряжений вводится через тензор напряжений Пиола $\mathbf{P} = \mathbf{P}(\overset{0}{\nabla} \mathbf{R})$ на основании определения элементарной работы деформации (δ – знак вариации) [4]:

$$\frac{\partial \vartheta}{\partial \overset{0}{\nabla} \mathbf{R}} = \mathbf{P} (\delta \vartheta = \mathbf{P} \cdot \delta \overset{0}{\nabla} \mathbf{R}^T). \quad (2)$$

С помощью соотношений $\mathbf{T} = (\sqrt{I_3})^{-1} \overset{0}{\nabla} \mathbf{R}^T \cdot \mathbf{P}$, $\frac{\partial \vartheta}{\partial \overset{0}{\nabla} \mathbf{R}} = 2 \frac{\partial \vartheta}{\partial \mathbf{G}} \cdot \overset{0}{\nabla} \mathbf{R}$, $\frac{\partial \vartheta}{\partial \mathbf{G}} = -\mathbf{G}^{-1} \cdot \frac{\partial \vartheta}{\partial \mathbf{G}^{-1}} \cdot \mathbf{G}^{-1}$

(I_3 – главный третий инвариант \mathbf{G} , $\mathbf{G} = \overset{0}{\nabla} \mathbf{R}^T \cdot \overset{0}{\nabla} \mathbf{R}$ – мера деформации Коши – Грина) потенциал записывается через тензор напряжений Коши [4].

Потенциал скоростей напряжений определяется через О-производную, которая для изотропного материала совпадает с производной Грина – Нахди и с производной Яуманна $\overset{w}{\mathbf{T}} = \mathbf{T} - \mathbf{W} \cdot \mathbf{T} + \mathbf{T} \cdot \mathbf{W}$ ($\mathbf{W} = 2^{-1}(\nabla \mathbf{v}^T - \nabla \mathbf{v})$ – тензор вихря):

$$\frac{\partial \varphi}{\partial \mathbf{D}} = \overset{\Omega}{\mathbf{T}} (\delta \varphi = \overset{\Omega}{\mathbf{T}} \cdot \delta \mathbf{D}). \quad (3)$$

Будем считать, что согласно (3) тензор $\overset{\Omega}{\mathbf{T}}$ имеет потенциал $\varphi = \varphi(\mathbf{D})$ или выражение $\overset{\Omega}{\mathbf{T}} \cdot \delta \mathbf{D}$ является потенциальным.

Уравнение состояния упругого тела в форме Фингера по (2) имеет вид [4]

$$\mathbf{T} = 2(\sqrt{I_3})^{-1}(\varphi_0 \mathbf{E} + \varphi_1 \mathbf{F} + \varphi_2 \mathbf{F}^2), \quad (4)$$

где φ_i – скалярные функции от I_1, I_2, I_3 (главных первого, второго и третьего инвариантов меры Фингера $\mathbf{F} = \overset{0}{\nabla} \mathbf{R}^T \cdot \overset{0}{\nabla} \mathbf{R}$); \mathbf{E} – единичный тензор;

$$\mathbf{T} = 2(\sqrt{I_3})^{-1} \overset{0}{\nabla} \mathbf{R}^T \cdot \frac{\partial \varphi}{\partial \mathbf{G}} \cdot \overset{0}{\nabla} \mathbf{R}; \quad \varphi_0 = I_3 \frac{\partial \varphi}{\partial I_3}; \quad \varphi_1 = \left(\frac{\partial \varphi}{\partial I_1} + I_1 \frac{\partial \varphi}{\partial I_2} \right); \quad \varphi_2 = -\frac{\partial \varphi}{\partial I_2}. \quad (5)$$

Если выполняется только соотношение (4), материал называют «упругим». Если существует такой скаляр $\varphi = \varphi(I_1, I_2, I_3)$ – удельная потенциальная энергия деформации (мера запасенной энергии), что выполняются дополнительно соотношения (5), то материал называют гиперупругим. Исключая потенциал φ из (5) ($\frac{\partial^2 \varphi}{\partial I_1 \partial I_2} = \frac{\partial^2 \varphi}{\partial I_2 \partial I_1}$, $\frac{\partial^2 \varphi}{\partial I_1 \partial I_3} = \frac{\partial^2 \varphi}{\partial I_3 \partial I_1}$, $\frac{\partial^2 \varphi}{\partial I_2 \partial I_3} = \frac{\partial^2 \varphi}{\partial I_3 \partial I_2}$), получаем условия интегрируемости (условия существования потенциала):

$$-\frac{\partial \varphi_2}{\partial I_1} = \frac{\partial}{\partial I_2} (\varphi_1 + I_1 \varphi_2), \quad \frac{\partial \varphi_0}{\partial I_1} = I_3 \frac{\partial}{\partial I_3} (\varphi_1 + I_1 \varphi_2), \quad \frac{\partial \varphi_0}{\partial I_2} = -I_3 \frac{\partial \varphi_2}{\partial I_3}. \quad (6)$$

Закон Мурнагана имеет вид [4, 5]

$$\begin{aligned} \mathbf{T} &= 2(\sqrt{I_3})^{-1}(\varphi_0 \mathbf{E} + \varphi_1 \mathbf{F} + \varphi_2 \mathbf{F}^2), \quad \varphi_0 = a_0 I_3, \quad \varphi_1 = b_0 + b_1 I_1 + b_2 I_1^2 + b_3 I_2, \quad \varphi_2 = c_0 + c_1 I_1, \\ a_0 &= 2^{-1} \nu_3, \quad b_0 = 16^{-1}(-12\lambda - 8\mu + 9\nu_1 + 18\nu_2 + 8\nu_3), \quad b_1 = 8^{-1}(2\lambda - 3\nu_1 - 4\nu_2), \\ b_2 &= 16^{-1}(\nu_1 + 2\nu_2), \quad b_3 = -4^{-1}(\nu_2 + 2\nu_3), \quad c_0 = 4^{-1}(2\mu - 3\nu_2 - 4\nu_3), \quad c_1 = -b_3, \end{aligned} \quad (7)$$

где λ, μ – постоянные Ляме второго и ν_1, ν_2, ν_3 третьего порядков соответственно.

Получим необходимые вспомогательные соотношения. Обозначим $d_i = \mathbf{E} \cdot \mathbf{F}^i$, $d'_i = \mathbf{E} \cdot (\text{dev } \mathbf{F})^i$, $i = 2, 3, \dots$. По определению второго инварианта ($I_2 = 2^{-1}(I_1^2 - d_2)$) и теореме Гамильтона – Кэли ($\mathbf{F}^3 = I_1 \mathbf{F}^2 - I_2 \mathbf{F} + I_3 \mathbf{E}$, $(\text{dev } \mathbf{F})^3 = 2^{-1} d'_2 \text{dev } \mathbf{F} + 3^{-1} d'_3 \mathbf{E}$) получаем

$$d_2 = I_1^2 - 2I_2, \quad d_3 = 3I_3 + I_1^3 - 3I_1 I_2, \quad d_4 = 4I_1 I_3 + I_1^4 - 3I_1^2 I_2 + 2I_2^2, \quad d'_4 = \frac{1}{2}(d'_2)^2; \quad (8)$$

$$d_2 = 3^{-1} I_1^2 + d'_2, \quad d_3 = 9^{-1} I_1^3 + I_1 d'_2 + d'_3, \quad d_4 = 27^{-1} I_1^4 + 3^{-1} 2I_1^2 d'_2 + 3^{-1} 4I_1 d'_3 + 2^{-1} (d'_2)^2; \quad (9)$$

$$I_2 = 3^{-1} I_1^2 - 2^{-1} d'_2, \quad I_3 = 27^{-1} I_1^3 - 6^{-1} I_1 d'_2 + 3^{-1} d'_3.$$

С учетом (8) и (9) находим

$$d_2 - 3^{-1}I_1^2 = d'_2, \quad d_3 - 3^{-1}I_1d_2 = 3^{-1}2I_1d'_2 + d'_3, \quad d_4 - 3^{-1}d_2^2 = 9^{-1}4I_1^2d'_2 + 3^{-1}4I_1d'_3 + 6^{-1}(d'_2)^2, \\ (d_4 - 3^{-1}d_2^2)(d_2 - 3^{-1}I_1^3) - (d_3 - 3^{-1}I_1d_2)^2 = 6^{-1}(d'_2)^3 - (d'_3)^2.$$

Поскольку $(d'_2)^3 \geq 6(d'_3)^2$ [4], при $d'_2 > 0$ из (9) получаем

$$d_2 - 3^{-1}I_1^2 > 0, \quad d_3 - 3^{-1}I_1d_2 > 0, \quad d_4 - 3^{-1}d_2^2 > 0, \quad (d_4 - 3^{-1}d_2^2)(d_2 - 3^{-1}I_1^2) \geq (d_3 - 3^{-1}I_1d_2)^2.$$

Если положить

$$r_2 = d_2 - 3^{-1}I_1^2, \quad r_1 = -(d_3 - 3^{-1}I_1d_2), \quad r_0 = d_4 - 3^{-1}d_2^2, \quad (10)$$

последнее неравенство запишется как

$$r_0r_2 \geq r_1^2. \quad (11)$$

С точностью до малых величин d'_i из (7) – (9) находим приближенные оценки

$$\varphi_1 = -2^{-1}\mu - \nu_3 + 4^{-1}(\lambda - \nu_2 - 4\nu_3)(I_1 - 3); \\ \varphi_2 = 2^{-1}(\mu + \nu_3) + 4^{-1}(\nu_2 + 2\nu_3)(I_1 - 3), \quad (12)$$

где $I_1 - 3$ – также малая величина.

3. Вектор нормали к поверхности девиаторного сечения для материала Мурнагана

Вычисляем из (7) девиатор \mathbf{Q} :

$$\dot{\mathbf{F}} = \nabla \mathbf{v}^T \cdot \mathbf{F} + \mathbf{F} \cdot \nabla \mathbf{v}, \quad (\mathbf{F}^2)^\dot{=} = \dot{\mathbf{F}} \cdot \mathbf{F} + \mathbf{F} \cdot \dot{\mathbf{F}} = \\ = \nabla \mathbf{v}^T \cdot \mathbf{F}^2 + \mathbf{F}^2 \cdot \nabla \mathbf{v} + \mathbf{F} \cdot (\nabla \mathbf{v}^T + \nabla \mathbf{v}) \cdot \mathbf{F}, \\ \dot{I}_1 = \frac{\partial I_1}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} = 2\mathbf{F} \cdot \mathbf{D}, \quad \dot{I}_2 = \frac{\partial I_2}{\partial \mathbf{F}} \cdot \dot{\mathbf{F}} = 2(I_1\mathbf{F} \cdot \mathbf{D} - \mathbf{F}^2 \cdot \mathbf{D}), \quad \dot{I}_3 = I_3\mathbf{F}^{-1} \cdot \dot{\mathbf{F}} = 2I_3\mathbf{E} \cdot \mathbf{D} = 0, \\ \dot{\varphi}_i = \frac{\partial \varphi_i}{\partial I_1} \cdot \dot{I}_1 + \frac{\partial \varphi_i}{\partial I_2} \cdot \dot{I}_2 = 2\frac{\partial \varphi_i}{\partial I_1} \mathbf{F} \cdot \mathbf{D} + 2\frac{\partial \varphi_i}{\partial I_2} (I_1\mathbf{F} \cdot \mathbf{D} - \mathbf{F}^2 \cdot \mathbf{D}) = 2A_i\mathbf{F} \cdot \mathbf{D} - 2B_i\mathbf{F}^2 \cdot \mathbf{D}, \quad (13) \\ \dot{\varphi}_i = 2A_i\mathbf{F} \cdot \mathbf{D} - 2B_i\mathbf{F}^2 \cdot \mathbf{D} \quad (A_i = \frac{\partial \varphi_i}{\partial I_1} + I_1 \frac{\partial \varphi_i}{\partial I_2}, B_i = \frac{\partial \varphi_i}{\partial I_2}).$$

Учитывая, что $\mathbf{E} \cdot \delta \mathbf{D} = 0$, девиатор \mathbf{Q} с точностью до шарового тензора можно представить в виде

$$\mathbf{Q} = 2(\sqrt{I_3})^{-1}(\dot{\varphi}_1\mathbf{F} + \dot{\varphi}_2\mathbf{F}^2 + \varphi_1(\mathbf{F} \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}) + \varphi_2(2\mathbf{F} \cdot \mathbf{D} \cdot \mathbf{F} + \mathbf{F}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}^2)). \quad (14)$$

По представлению (14) получаем, что девиатор \mathbf{Q} имеет потенциал

$$\mathbf{Q} = \frac{\partial}{\partial \mathbf{D}} (2(\sqrt{I_3})^{-1}(A_1(\mathbf{F} \cdot \mathbf{D})^2 - 2B_1\mathbf{F} \cdot \mathbf{D}\mathbf{F}^2 \cdot \mathbf{D} - 2B_2(\mathbf{F}^2 \cdot \mathbf{D})^2 + \varphi_1\mathbf{F} \cdot \mathbf{D}^2 + \varphi_2(2\mathbf{F} \cdot \mathbf{D} \cdot \mathbf{F} \cdot \mathbf{D} + \mathbf{F}^2 \cdot \mathbf{D}^2))).$$

Вектор нормали к поверхности девиаторного сечения запишем двумя возможными способами:

$$\mathbf{N} = (\text{dev } \mathbf{T})^2 - z \text{dev } \mathbf{T} + 3^{-1} 2J_2 \mathbf{E} \quad (J_2 = -2^{-1} \mathbf{E} \cdot (\text{dev } \mathbf{T})^2) \quad (15)$$

и, как будет удобно для доказательства последующих утверждений,

$$\mathbf{N} = q\mathbf{F}^2 + \mathbf{F} - 3^{-1}(qd_2 + I_1)\mathbf{E}. \quad (16)$$

Обозначим девиатор $\delta\mathbf{D}$ как вариацию девиатора \mathbf{D} . Так как выражение $\mathbf{Q} \cdot \delta\mathbf{D}$ является потенциальным, то должно быть потенциальным выражение $\mathbf{Q} \cdot \mathbf{N} \mathbf{N} \cdot \delta\mathbf{D}$. В последнем выражении будет только один чисто непотенциальный член $\gamma(\mathbf{F}^2 \cdot \mathbf{D} \mathbf{F} \cdot \delta\mathbf{D} - \mathbf{F} \cdot \mathbf{D} \mathbf{F}^2 \cdot \delta\mathbf{D})$. Следовательно, выполняется $\gamma = 0$ по следствию из уравнения (1).

В соотношении (14) опустим множитель $4(\sqrt{I_3})^{-1}$. Используя теорему Гамильтона – Кэли, соотношения (7) – (10), (13), (16) и учитывая условие несжимаемости $\mathbf{E} \cdot \mathbf{D} = 0$, имеем

$$\begin{aligned} \mathbf{N} \cdot \delta\mathbf{D} &= q\mathbf{F}^2 \cdot \delta\mathbf{D} + \mathbf{F} \cdot \delta\mathbf{D} \quad (\mathbf{E} \cdot \delta\mathbf{D} = 0), \\ \mathbf{Q} \cdot \mathbf{N} &= 2^{-1}(\dot{\varphi}_1 \mathbf{F} + \dot{\varphi}_2 \mathbf{F}^2) \cdot \mathbf{N} + 2^{-1} \varphi_1 (\mathbf{F} \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}) \cdot \mathbf{N} + 2^{-1} \varphi_2 (2\mathbf{F} \cdot \mathbf{D} \cdot \mathbf{F} + \mathbf{F}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}^2) \cdot \mathbf{N}, \\ 2^{-1}(\dot{\varphi}_1 \mathbf{F} + \dot{\varphi}_2 \mathbf{F}^2) \cdot \mathbf{N} &= 2^{-1}(\dot{\varphi}_1 \mathbf{F} \cdot \mathbf{N} + \dot{\varphi}_2 \mathbf{F}^2 \cdot \mathbf{N}) = \\ &= ((-B_1 \mathbf{F} \cdot \mathbf{N} - B_2 \mathbf{F}^2 \cdot \mathbf{N}) \mathbf{F}^2 + (A_1 \mathbf{F} \cdot \mathbf{N} + A_2 \mathbf{F}^2 \cdot \mathbf{N}) \mathbf{F}) \cdot \mathbf{D}, \\ 2^{-1} \varphi_1 (\mathbf{F} \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}) \cdot \mathbf{N} &= \varphi_1 \mathbf{F} \cdot \mathbf{N} \cdot \mathbf{D} = \varphi_1 (q\mathbf{F}^3 + \mathbf{F}^2 - 3^{-1}(qd_2 + I_1)\mathbf{F}) \cdot \mathbf{D} = \\ &= \varphi_1 (\xi_2 \mathbf{F}^2 + \xi_1 \mathbf{F}) \cdot \mathbf{D}, \\ 2^{-1} \varphi_2 (2\mathbf{F} \cdot \mathbf{D} \cdot \mathbf{F} + \mathbf{F}^2 \cdot \mathbf{D} + \mathbf{D} \cdot \mathbf{F}^2) \cdot \mathbf{N} &= 2\varphi_2 \mathbf{F}^2 \cdot \mathbf{N} \cdot \mathbf{D} = 2\varphi_2 (q\mathbf{F}^4 + \mathbf{F}^3 - 3^{-1}(qd_2 + I_1)\mathbf{F}^2) \cdot \mathbf{D} = \\ &= 2\varphi_2 (\eta_2 \mathbf{F}^2 + \eta_1 \mathbf{F}) \cdot \mathbf{D}, \\ \mathbf{N} \cdot \mathbf{F} &= r_2 - qr_1, \xi_2 = 1 + qI_1, \xi_1 = -3^{-1}I_1 - q(2^{-1}I_1^2 - 6^{-1}d_2), \\ \mathbf{N} \cdot \mathbf{F}^2 &= -r_1 + qr_0, \eta_2 = 3^{-1}2I_1 + q(2^{-1}I_1^2 + 6^{-1}d_2), \eta_1 = -I_2 + q(I_3 - I_1 I_2). \end{aligned} \quad (17)$$

Вычисляя $\mathbf{Q} \cdot \mathbf{N} \mathbf{N} \cdot \delta\mathbf{D}$ с точностью до потенциальных членов, находим

$$\begin{aligned} \gamma &= \gamma_0 + \gamma_1 \varphi_1 + 2\varphi_2 \gamma_2, \gamma_2 = \eta_2 - q\eta_1, \gamma_1 = \xi_2 - q\xi_1; \\ \gamma_0 &= -\mathbf{F} \cdot \mathbf{N} B_1 - \mathbf{F}^2 \cdot \mathbf{N} B_2 - q(\mathbf{F} \cdot \mathbf{N} A_1 + \mathbf{F}^2 \cdot \mathbf{N} A_2). \end{aligned} \quad (18)$$

Из (17), (18) для $j = 0, 1, 2$ получаем

$$0 = \gamma = \gamma_0 + \gamma_1 \varphi_1 + 2\varphi_2 \gamma_2, \quad \gamma_j = q^2 Q_{2j} + q Q_{1j} + Q_{0j}; \quad (19)$$

$$0 = \gamma = Q_2 q^2 + Q_1 q + Q_0, \quad Q_j = Q_{j0} + \varphi_1 Q_{j1} + 2\varphi_2 Q_{j2}; \quad (20)$$

$$\begin{aligned} Q_{20} &= r_1 A_1 - r_0 A_2, & Q_{10} &= -r_2 A_1 + r_1 (A_2 + B_1) - r_0 B_2, & Q_{00} &= -r_2 B_1 + r_1 B_2, \\ Q_{21} &= 2^{-1} I_1^2 - 6^{-1} d_2, & Q_{11} &= 3^{-1} 4 I_1, & Q_{01} &= 1, \\ Q_{22} &= -(I_3 - I_1 I_2), & Q_{12} &= I_2 + 2^{-1} I_1^2 + 6^{-1} d_2, & Q_{02} &= 3^{-1} 2 I_1. \end{aligned} \quad (21)$$

Используя (7) – (9), (12), (20) с точностью до величин $I_1 - 3$ и d'_i более высоких порядков, получаем

$$\begin{aligned} Q_0 &= 2^{-1}3\mu + v_3 + (I_1 - 3)(4^{-1}\lambda + 3^{-1}2\mu + 4^{-1}3v_2 + 3^{-1}5v_3) + (-4^{-1}v_2 - 2^{-1}v_3)d'_2; \\ Q_1 &= -\mu - 3v_3 + (I_1 - 3)(\lambda + 3^{-1}\mu - 2^{-1}v_2 - 4v_3) + (3^{-1}\mu + 6^{-1}v_2 + 2^{-1}v_3)d'_2; \\ Q_2 &= -\mu - 3v_3 + (I_1 - 3)(\lambda + 3^{-1}\mu - 2^{-1}v_2 - 3^{-1}14v_3) - (12^{-1}11\mu + 3^{-1}2v_2 + 6^{-1}11v_3)d'_2, \end{aligned} \quad (22)$$

где $(I_1 - 3)$ и d'_2 – малые величины, так что значения Q_i определяются главным образом первыми слагаемыми.

Если $Q_2 \neq 0$, всегда имеются два, возможно, комплексных или равных действительных значения q_1, q_2 решения уравнения (20) $Q_2q^2 + Q_1q + Q_0 = 0$. Им соответствуют два девиатора $\mathbf{N}_i = q_i\mathbf{F}^2 + \mathbf{F} - 3^{-1}(q_id_2 + I_1)\mathbf{E}$ ($i = 1, 2$). Условие их ортогональности $\mathbf{N}_1 \cdot \mathbf{N}_2 = 0$ имеет вид $r_0q_1q_2 - r_1(q_1 + q_2) + r_2 = 0$ по (10), (16), и с учетом (20) находим

$$r_2Q_2 + r_1Q_1 + r_0Q_0 = 0, \quad (23)$$

где величины r_i, Q_i определяются по (10), (20) и (21).

Значения q_1, q_2 будут вещественными, если выполняется $Q_1^2 - 4Q_0Q_2 > 0$ при $Q_2 \neq 0$. Недопустимыми значениями упругих постоянных согласно (22) являются значения $\{(\mu, v_3), (v_3 \leq -3^{-1}\mu) \wedge (v_3 \geq -\mu)\}$. При $Q_0Q_2 > 0$ возможно следующее рассуждение. Используя (11), получаем $2Q_0Q_2(2r_1^2 - r_0r_2) \leq 2Q_0Q_2r_0r_2 < Q_0^2r_0^2 + Q_2^2r_2^2$, $4Q_0Q_2 < r_1^{-2}(r_0Q_0 + r_2Q_2)^2 = Q_1^2$ по (23).

Докажем теперь условие ортогональности (23). Непосредственно проверяем по (8) – (10), (21), (23) тождественность равенств

$$r_2Q_{2i} + r_1Q_{1i} + r_0Q_{0i} = 0 \quad (i = 1, 2). \quad (24)$$

Далее вычисляем

$$r_2Q_{20} + r_1Q_{10} + r_0Q_{00} = -(r_0r_2 - r_1^2)(A_2 + B_1) = 0 \quad (25)$$

согласно одному из условий интегрируемости (6): $\frac{\partial\varphi_2}{\partial I_1} + I_1 \frac{\partial\varphi_2}{\partial I_2} = -\frac{\partial\varphi_1}{\partial I_2}$.

Умножая первое уравнение (24) на φ_1 , второе на $2\varphi_2$ и складывая их с уравнением (25), находим $0 = \varphi_1(r_2Q_{21} + r_1Q_{11} + r_0Q_{01}) + 2\varphi_2(r_2Q_{22} + r_1Q_{12} + r_0Q_{02}) + r_2Q_{20} + r_1Q_{10} + r_0Q_{00} = r_2(Q_{20} + \varphi_1Q_{21} + 2\varphi_2Q_{22}) + r_1(Q_{10} + \varphi_1Q_{11} + 2\varphi_2Q_{12}) + r_0(Q_{00} + \varphi_1Q_{01} + 2\varphi_2Q_{02}) = r_2Q_2 + r_1Q_1 + r_0Q_0$.

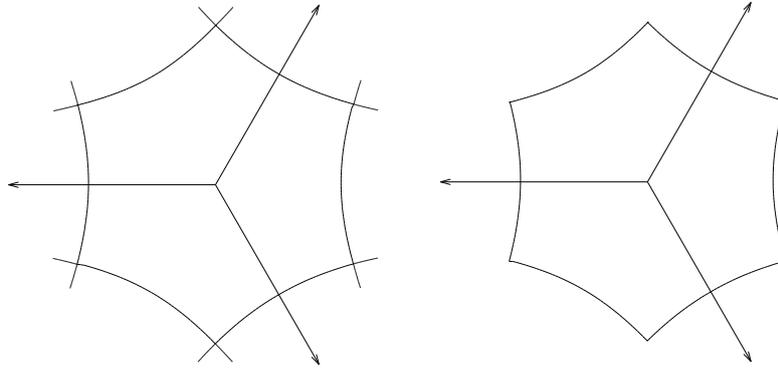
Следовательно, условие ортогональности (23) выполняется вследствие гиперупругости материала Мурнагана. К такому же выводу приходим в [15] для закона Синьорини. Можно показать, что этот факт имеет место при использовании в данном подходе любого гиперупругого материала. Таким образом, для этого материала при изотропии оба способа определения девиаторов $\mathbf{N}_1, \mathbf{N}_2$ приводят к одному результату. Для негиперупругого материала возможно появление новых решений при вычислении $\gamma = 0$, поскольку при использовании собственных векторов симметричного оператора таких решений не получается.

Отметим, что для практического использования девиатора \mathbf{N} больше подходит представление (15). Его следует дополнить с учетом (20) и (21) соотношениями, которые дадим без вывода (аналогичные выкладки приведены в [16]):

$$\begin{aligned} Z_2 z^2 + Z_1 z + Z_0 &= 0, Z_0 = Q_0 \beta_4^2 + Q_1 \beta_2 \beta_4 + Q_2 \beta_2^2; \\ Z_1 &= 2Q_0 \beta_3 \beta_4 + Q_1 (\beta_1 \beta_4 + \beta_2 \beta_3) + 2Q_2 \beta_1 \beta_2, Z_2 = Q_0 \beta_3^2 + Q_1 \beta_1 \beta_3 + Q_2 \beta_1^2; \\ \beta_1 &= -2\varphi_2, \beta_2 = 4(\sqrt{I_3})^{-1} (\varphi_2^2 (I_1^2 - I_2) + 2\varphi_1 \varphi_2 I_1 + 2\varphi_2 \varphi_3 + \varphi_1^2); \\ \beta_3 &= -\varphi_1, \beta_4 = 4(\sqrt{I_3})^{-1} (\varphi_2^2 (I_3 - I_1 I_2) - 2\varphi_1 \varphi_2 I_2 + 2\varphi_1 \varphi_3) \quad (\varphi_3 = -3^{-1} (\varphi_1 I_1 + \varphi_2 d_2)). \end{aligned}$$

Представления (15) и (16) следует, конечно, нормировать и задать z, q с учетом знака $J_3 = 3^{-1} \mathbf{E} \cdot (\text{dev } \mathbf{T})^3$:

$$z = (-Z_1 + s\sqrt{Z_1^2 - 4Z_2 Z_0}) (2Z_2)^{-1}, \quad q = (-Q_1 + s\sqrt{Q_1^2 - 4Q_2 Q_0}) (2Q_2)^{-1}, \quad s = \begin{cases} 1 (J_3 \geq 0), \\ -1 (J_3 \leq 0). \end{cases} \quad (26)$$



Образование кривой пластичности

На рисунке в плоскости $\{\mathbf{W}_1, \mathbf{W}_2\}$ представлены результаты расчета кривой пластичности. Стрелками обозначены проекции на эту плоскость базисных диад $\mathbf{e}_i \mathbf{e}_i$. Данные о постоянных Ляме взяты для молибдена из [4]: $\lambda = 1,63, \mu = 1,37, \nu_1 = -4,29, \nu_2 = -2,58, \nu_3 = -3,67 (10^5 \text{ МПа})$. Слева показаны представители двух семейств кривых, из которых получается изображенная справа кривая.

Заключение

Для вектора нормали к поверхности девиаторного сечения предлагается способ его выбора в условиях любой анизотропии из пяти вычисленных собственных векторов оператора \mathbf{Q} . Из условия потенциальности девиатора $\mathbf{Q} \cdot \mathbf{N} \mathbf{N}$ по (1) находится в явном виде вектор \mathbf{N}_0 при изотропии согласно (16), (20), (21), (26). Затем он сравнивается с вычисленными собственными векторами. Ближайший из них к вектору \mathbf{N}_0 будет искомым вектором нормали при анизотропии. Способ выбора эффективен вне окрестности значения $J_3 = 0$. В самой окрестности требуется дополнительная процедура визуализации возможных векторов нормалей. Условие, что вектор нормали является одним из собственных векторов оператора \mathbf{Q} для обобщенно упругого материала, является только достаточным. Однако, например, для тела Сетха [4] оно является и необходимым.

Список литературы

1. Жилин, П.А. Основные уравнения неупругих сред / П.А. Жилин // Актуальные проблемы механики. Труды XXVIII летней школы. – СПб., 2001. – С. 14–58.
2. Швед, О.Л. О возможных определяющих соотношениях нелинейной упругопластичности / О.Л. Швед // Труды VII Всерос. конф. по механике деформируемого твердого тела. – Ростов н/Д., 2013. – Т. II. – С. 219–223.
3. Швед, О.Л. Определяющие соотношения ортотропного упругопластического материала / О.Л. Швед // Там же. – С. 224–228.
4. Лурье, А.И. Нелинейная теория упругости / А.И. Лурье. – М. : Наука, 1980. – 512 с.
5. Murnagan, F.D. Finite deformation of an elastic solid / F.D. Murnagan. – N.Y. : Dover, 1967. – 140 p.
6. Naghdi, P.M. A critical review of the state of finite plasticity / P.M. Naghdi // ZAMP. – 1990. – Vol. 41, № 3. – P. 315–394.
7. Ключников, В.Д. Физико-математические основы прочности и пластичности / В.Д. Ключников. – М. : МГУ, 1994. – 189 с.
8. Белл, Дж.Ф. Экспериментальные основы механики деформируемых твердых тел. Ч. II. Конечные деформации / Дж.Ф. Белл. – М. : Наука, 1984. – 432 с.
9. Поздеев, А.А. Большие упругопластические деформации: теория, алгоритмы, приложения / А.А. Поздеев, П.В. Трусов, Ю.И. Няшин. – М. : Наука, 1986. – 232 с.
10. Швед, О.Л. Математическое моделирование процесса прямого выдавливания свинца / О.Л. Швед, А.А. Абрамов // Информатика. – 2007. – № 4 (16). – С. 133–136.
11. Швед, О.Л. Исследование эффекта двойного бочкообразования при обработке металлов давлением с использованием системы LS-DYNA / О.Л. Швед, А.А. Абрамов // Информатика. – 2009. – № 1 (21). – С. 17–24.
12. Левитас, В.И. Большие упругопластические деформации материалов при высоком давлении / В.И. Левитас. – Киев : Наукова думка, 1987. – 232 с.
13. Бриджмен, П.В. Исследование больших пластических деформаций и разрыва / П.В. Бриджмен. – М. : Изд-во иностр. лит., 1955. – 444 с.
14. Швед, О.Л. Определение тензора упругого спина в нелинейной теории пластичности / О.Л. Швед // Весці НАН Беларусі. Сер. фіз.-мат. навук. – 2009. – № 1. – С. 52–58.
15. Махнач, В.И. Начальное условие пластичности при конечных деформациях / В.И. Махнач, О.Л. Швед // Вести НАН Беларусі. Сер. фіз.-мат. навук. – 2003. – № 3. – С.95–99.
16. Швед, О.Л. Вариант определяющих соотношений идеальнопластической среды при конечных деформациях / О.Л. Швед // Весці НАН Беларусі. Сер. фіз.-тэх. навук. – 2001. – № 1. – С. 120–125.

Поступила 6.02.2014

*Объединенный институт проблем информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: swed@newman.bas-net.by*

O.L. Shwed

**DEFINING DEVIATORIC SECTION OF THE FLUIDITY SURFACE
FOR MATHEMATICAL MODELING OF ELASTOPLASTIC
BEHAVIOR OF MATERIALS**

A prospective approach to the development of nonlinear models of elastoplasticity based on the generalization of nonlinear elasticity models is considered. For generalized Murnagan information, a way of finding a normal vector to the surface of deviatoric section under the condition of anisotropy is suggested.

УДК 681.3.06

А.Ю. Переварюха

РАЗНОВИДНОСТИ АПЕРИОДИЧЕСКОЙ ДИНАМИКИ В СОБЫТИЙНО-УПРАВЛЯЕМОЙ ПОПУЛЯЦИОННОЙ МОДЕЛИ

Предлагается популяционная модель событийно-стадийной структуры, включающая непрерывные и дискретные составляющие. Динамика гибридной системы анализируется в вычислительной среде на основе численного решения последовательности задач Коши для системы дифференциальных уравнений убыли поколений. Рассматривается динамика функциональной итерации, которая имеет два локальных экстремума и характеризует непостоянство эффективности воспроизводства рыб. Устанавливается появление переходного аperiodического режима при возможности притяжения траектории к двум аттракторам. После бифуркации исчезновения двух нетривиальных стационарных точек возникает интервальный аттрактор, для которого наблюдается явление граничного кризиса.

Введение

Математические методы моделирования биологических процессов продолжают развиваться в нескольких отдельных междисциплинарных направлениях. Изначально предложения по формализации популяционной динамики заключались в разработке систем дифференциальных уравнений с правыми частями различного вида, отражающими некоторые теоретические предположения о характере взаимодействия биомассы конкурирующих видов.

В задачах моделирования сообществ гидробионтов в 1970 гг. развивались алгоритмические методы, основанные на исследовании дискретных вычислительных структур с применением ЭВМ. Советской школой предложено значительное число разнообразных дискретных моделей для промысловой ихтиологии: В.В. Меншуткина, А.Б. Казанского, В.В. Суханова и др. Развитием подхода стало создание В.В. Иванищевым и В.В. Михайловым в 1984 г. высокоуровневого языка алгоритмических сетей, предназначенного для описания процесса пользователем в виде ориентированного «функционального» графа операторов. Возможности модельного описания и прогнозирования популяционных процессов столкнулись с фундаментальными проблемами теории универсальности поведения нелинейных систем [1]. Дискретно-матричные популяционные модели, ориентированные на вычислительные методы исследования, обладают нетривиальными возможностями изменения поведения с хаотическими и циклическими режимами.

В настоящей статье на основе представлений об экологических особенностях воспроизводства осетровых рыб предложена гибридная (событийно-управляемая) модель динамики численности, обладающая свойством трансформации двух видов аperiodического поведения траектории.

1. Постановка задачи популяционного моделирования

В основе моделей эксплуатируемых популяций рыб лежит формализация баланса воспроизводства и смертности от различных факторов. Естественная убыль на ранних этапах жизни у крупных анадромных рыб очень велика, и ее изменения критически сказываются на благополучии популяции. Опыт наблюдений показал, что среднее пополнение R от величины родительского запаса S редко удовлетворительно описывается линейным или кусочно-линейным соотношением $R = f(S) = aS, S < K$ $f(S) = X = cost, S > K$.

Пополнением будем считать численность поколения от одного нереста, дожившего до установленного момента. Для разных видов рыб этот момент может определяться достижением промысловых размеров, прохождением периода адаптации к морскому периоду жизни или окончанием полового созревания.

В ихтиологии существует достаточно развитое направление исследований, объясняющее закономерности изменения эффективности воспроизводства. Основной его целью является

определение зависимости для прогнозирования скорости восполнения промысловых запасов на основе данных наблюдений. Обобщенная задача представляется противоречивой, так как очевидны различия экологических особенностей нереста разных рыб. Дискуссия о роли зависимости и конкретных функциональных свойствах имеет длительную историю среди биологов.

Автором обосновано предположение, что механизмы, определяющие зависимость, действуют для рыб с особыми условиями нерестового цикла. Такие рыбы (называемые анадромными) заходят при необходимой температуре в реки и нерестятся на ограниченных по площади пригодных русловых нерестилищах, как крупные лососевые и осетровые. Известно, что при повышенной плотности икры на грунте наблюдается ряд негативных для выживаемости явлений, связанных с гипоксией и токсикозом. Ограниченность пригодных нерестовых участков, например, характерна для волжской севрюги. В период масштабного гидростроительства возводились искусственные нерестилища. Для размножающихся в толще воды рыб успех репродуктивного процесса определяет благоприятное сочетание случайных флуктуаций условий среды, что снижает возможности применения детерминированных моделей.

Концепция моделей воспроизводства заключается в описании лимитирующих факторов $\nu(S)$, действующих на нерестилищах при повышенной плотности запаса с репродуктивным потенциалом a : $R = aS / \nu(S)$. Основоположником научного направления У. Рикером предложена экспоненциальная форма $\nu(S) = e^{bS}$, где b – показатель действия лимитирующих факторов. Дж. Шепард применил зависимость по аналогии с моделью Ферхюльста вида $\nu(S) = 1 + (S / K)^b$, где учтена критическая биомасса K . Анализ моделей проводился в виде зависящих от параметра функциональных итераций $x_{n+1} = f(x_n; a)$ в современной вычислительной среде моделирования.

Для траектории итераций функции Рикера при возрастании a характерно изменение поведения от устойчивого равновесия R^* к хаосу через известный каскад удвоений Фейгенбаума [2]. Аналогично бифуркации удвоения периода цикла наблюдаются для модели Шепарда. Ранее было показано, что бифуркационные параметры в двух моделях имеют противоположный смысл. Это влечет проблему биологической интерпретации результатов моделирования [3].

Аттрактор, возникающий в результате накопления каскада бифуркаций удвоения, является аналогом канторовского множества – замкнутого множества, не содержащего как внутренних, так и изолированных точек. Структурно хаотический аттрактор представляет собой результат объединения все уменьшающихся субинтервалов, которые составляют точки отрезка за исключением несчетного числа неустойчивых точек всех периодов 2^n и их прообразов. Анализ образования и свойств канторовских множеств является отдельной задачей при исследовании нелинейных дискретных моделей. Теория универсальности изменения поведения отображений, удовлетворяющих критериям теоремы Д. Синжера [4], описана достаточно подробно.

Выявление свойства хаотичности важно для оценки адекватности биологических моделей. Обычно для определения хаотичности используется свойство чувствительной зависимости от начальных условий, но на основе работы [5] можно ввести критерий хаотичности отображения отрезка $f: I \rightarrow I$ на основе топологической транзитивности: если для всех открытых подмножеств $U \subseteq V$, $U \neq \emptyset$ и $Y \subseteq V$, $Y \neq \emptyset$, $\exists n \geq 0$, выполняется $f^n(U) \cap Y \neq \emptyset$, то поведение хаотично.

Помимо хаотизации и соответственно эффекта экспоненциального разбегания близких траекторий в отображениях возможны и другие нелинейные эффекты, связанные с окнами периодичности. Отметим, что нелинейные эффекты в динамике делают проблематичной существенную интерпретацию поведения дискретных популяционных моделей, в особенности не относящихся к SU -семейству (унимодальных отображений со всюду отрицательным шварцианом).

Сведения о воспроизводстве севрюги и горбуши показали, что наблюдается выраженная неунимодальная зависимость с двумя высокими диапазонами эффективности воспроизводства, между которыми существует промежуток численности запаса, при котором для популяции характерна низкая способность к восстановлению. С биологической точки зрения можно обосно-

вать предположение, что характер зависимости является следствием различия факторов смертности на разных этапах развития молоди рыб.

2. Описание новой модели

Физиологи выделяют стадии развития молоди по мере формирования у них органов и характера передвижения. Изменения происходят по мере развития за счет питания, скорость которого представим в виде уравнения обратно пропорциональной плотности:

$$\frac{dw}{dt} = \frac{g}{N^k(t) + l}, k < 1, \quad (1)$$

где g – параметр объема доступных кормовых ресурсов, l учитывает ограничение скорости развития, не связанное с плотностью особей поколения. Важнейшие изменения, переход на активное питание и начало самостоятельной миграции можно считать событиями в динамике поколения, что позволит выделить стадии D_1, D_2, D_3 . Имеются основания для применения гибридной модели с изменяемой структурой при достижении некоторых выделяемых предикатами событий.

Предложение по формализации процесса формирования пополнения заключается в описании убыли начальной численности поколения дифференциальным уравнением на промежутке времени интервала уязвимости $[0, T]$, которое позволило бы в явном виде учесть различные факторы смертности и изменение данных факторов по мере развития особей:

$$\frac{dN}{dt} = \begin{cases} -(\alpha w(t)N(t) + U\beta)N(t), t < \tau; \\ -(\alpha_1 N(\tau) / w(\tau) + \beta)N(t), t > \tau, w(t) < w_{D_2}; \\ -\alpha_2 w(t)N^2(t), w(t) < w_{D_3}, \end{cases} \quad (2)$$

где α – коэффициент зависящей от плотности особей компенсационной смертности; β – коэффициент нейтральной убыли. От абстрактного «репродуктивного потенциала» разумно перейти к естественному показателю средней плодовитости λ , оцениваемой по данным мониторинга, так как у осетровых нет половых хромосом и дифференциация происходит эпигамно. Начальные условия для уравнений (1), (2): $w(0) = w_0$, $N(0) = \lambda S$. Время τ – длительность первой стадии с эндогенным питанием, для севрюги в среднем составляет восемь суток; w_D – условный уровень развития, при достижении которого меняется действие факторов смертности, что интерпретируется экологией обитания молоди при начале самостоятельной миграции. Логично, что интервал уязвимости не является постоянным и может растягиваться при замедлении скорости развития.

По данным о воспроизводстве волжской севрюги выявлено действие отрицательного эффекта группы (известного в литературе как *allege effect* [6]), когда при низкой плотности особей уменьшается вероятность встреч на нерестилищах, что сильно сокращает продуктивность нереста. Потому в правую часть (2) для D_1 вводится функция $U(S)$, которая быстро стремится к единице: $E(U) = [2, 1)$, так как эффект не может проявляться при исторически оптимальной для промысловой популяции численности запаса:

$$U(S) = 1 + \exp(-cS^2), \quad (3)$$

где параметр $c < 1$ определяет степень выраженности эффекта. Обыкновенное дифференциальное уравнение (ОДУ) (1) и набор правых частей (2) и (3) формируют непрерывно-дискретную вычислительную структуру.

Эффект Олли предложил учитывать А.Д. Базыкин, дополняя известное уравнение Ферхюльста для популяции $N(t)$ с репродуктивным коэффициентом r сомножителем $(N - m)$:

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right) (N - m);$$

когда численность меньше критической $N < m$, популяция вымирает. Способ используется и в современных работах [7]. Интересный подход предложили в 2013 г. авторы [8] в диффузионной модели «хищник – жертва» с нелинейной репродуктивной функцией жертвы:

$$F(N) = rN \left(1 - \frac{N}{K} - \frac{m}{N+d} \right).$$

При соотношении коэффициентов $m < d$ эффект Олли выражен слабо, соотношение $m > d$ подразумевает сильное его проявление. Однако в обоих методах действие проявляется для всех значений N .

3. Принципы алгоритмической реализации гибридной модели

Особенность дискретно-событийного подхода составляют переходы между состояниями моделируемой системы согласно графу всех возможных состояний. В применяемом методе на основе формализма гибридного автомата переключение реализуется между режимами изменения состояния. Режимам изменения сопоставлен набор форм правой части системы ОДУ из (2), алгоритм контроля предикатов определяет выбор решаемой в данный момент задачи Коши с инициализацией новых начальных условий. Промежуток интервала уязвимости разбит на последовательность кадров гибридного времени в инструментальной вычислительной среде AnyLogic. Алгоритмическое представление модели реализуется на основе автомата с таймированными и предикативными переходами (рис. 1). Множество решений задач Коши для допустимых начальных условий $S \in Z^+$ на интервале $t \in [0, T]$ определит зависимость, называемую в работах ихтиологов кривой воспроизводства популяции.

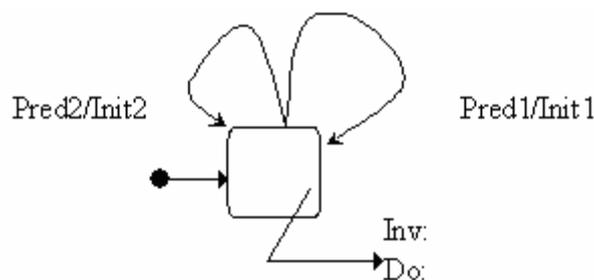


Рис. 1. Гибридный автомат системы с двумя событийными переходами

Каждому из двух переходов сопоставлен набор предикатов Pred1 или Pred2. За подтверждением истинности предикатов следует инициализация вычислительной задачи, для чего формируются переопределяемые начальные условия Init.

4. Вычислительный анализ свойств гибридной модели

В вычислительной среде получена неунимодальная «волнообразная» зависимость $R = N(T) = \varphi(S)$ запаса и пополнения, о причинах возникновения которой для крупных рыб писал Рикер в статье [9]. Зависимость (рис. 2) без учета действия промысловой смертности характеризуется четырьмя нетривиальными стационарными точками R_i^* , $i = 1, \dots, 4$, пересечениями кривой с биссектрисой координатного угла $R = N(T)$.

График второй итерации $\varphi(\varphi(S)) \equiv \varphi^2(S)$ показал устойчивость четвертой точки наряду с устойчивостью $R = 0$. Для дискретной динамической системы $R_{n+1} = \varphi(R_n)$ возможны качественно различные варианты поведения в зависимости от вычисленного значения функции в двух

точках локальных экстремумов $\min \varphi(R)$, $\max \varphi(R)$, $R_1^* < R < R_3^*$. Наибольший практический интерес представляет выполнение условий $\varphi(R_{\max} \pm \varepsilon) > R_3^*$, $\varphi(R_{\min} \pm \varepsilon) < R_1^*$. При выполнении условий в вычислительных экспериментах фиксируется образование в фазовом пространстве объекта, относящегося к разновидности непритягивающих хаотических множеств.



Рис. 2. Функциональная зависимость $\varphi(R)$ с четырьмя стационарными точками

Если для динамической системы существуют два аттрактора, то при исследовании необходимо определить границу их областей притяжения. В простейшем случае границей является неустойчивая «репеллерная» точка. В рассматриваемом случае границу составляет все множество прообразов неустойчивых точек R_i^* . Обе области притяжения в локальном диапазоне $[R_1^*, R_3^*]$ представляются несвязным объединением малых интервалов.

Канторовская структура границы приводит к появлению длительного переходного хаотического режима, реализующегося до момента $\varphi^z(R_0) > R_3^*$ (или $\varphi^z(R_0) < R_1^*$), достижение которого означает стремительное развитие неожиданной «вспышки» численности (рис. 3) популяции, что наблюдалось для горбуши. Число итераций z пребывания траектории в переходном аperiодическом режиме чувствительно зависит от выбора начальных условий [10].

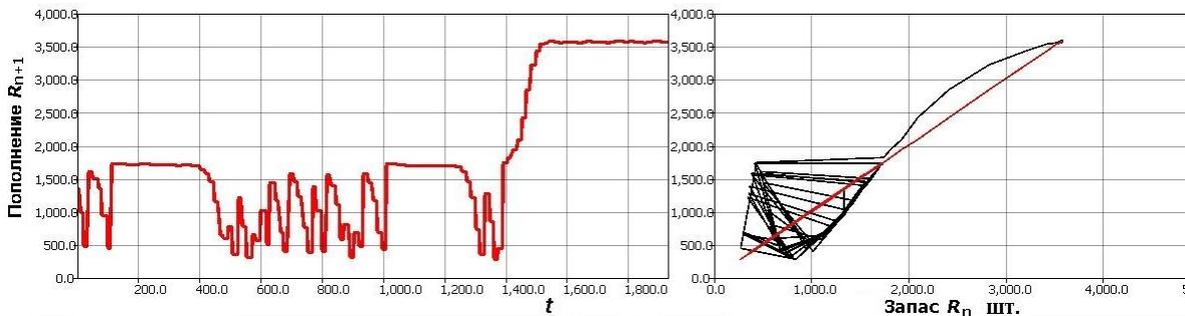


Рис. 3. Выход из режима переходного хаоса: временная и фазовая диаграммы

При рассмотрении в модели увеличения промышленной смертности изменяется конфигурация стационарных точек. Для динамической системы возможна обратная касательная бифуркация: слияние R_3^*, R_4^* с исчезновением стационарной точки при сохранении оставшихся R_1^*, R_2^* . В таком случае возможны два варианта, определенные смещающимся значением $\varphi_1(R_{\min})$ в точке минимума измененной зависимости (рис. 4). При выполнении условия $\varphi_1(R_{\min}) > R_1^*$ траектории притягиваются к интервальному аттрактору, неустойчивая точка R_1^* служит границей с областью притяжения тривиального равновесия (рис. 5). Тогда после обратной касательной бифуркации траектория моментально переходит к устойчивому аperiодическому режиму, что соответствует колебаниям в диапазоне низкой численности популяции без возможности восстановления. Подобная ситуация из-за последствий длительного перелома наблюдается сейчас с осетровыми Каспия, где промысел не был остановлен своевременно.

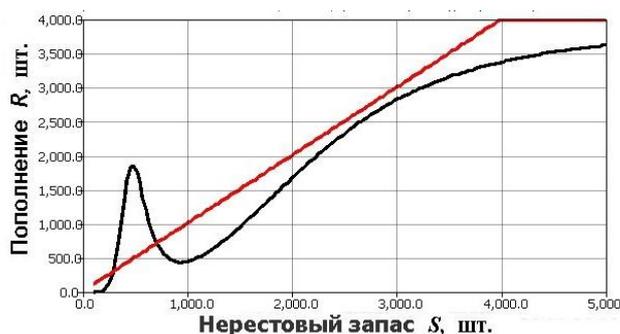


Рис. 4. Функциональная зависимость $\varphi(R)$ после обратной касательной бифуркации

В момент достижения $\varphi_1(R_{\min}) < R_1^*$ происходит граничный кризис интервального аттрактора [11]. При подобном кризисе аттрактор соприкасается с границей, теряет свойство инвариантности при сохранении локально несвязной структуры у вновь появившегося непритягивающего хаотического множества типа «хаотический репеллер» по классификации Гребоджи [12]. Единственным аттрактором остается тривиальное равновесие, что описывает неминуемую деградацию популяции после короткого переходного аперiodического режима флуктуаций.

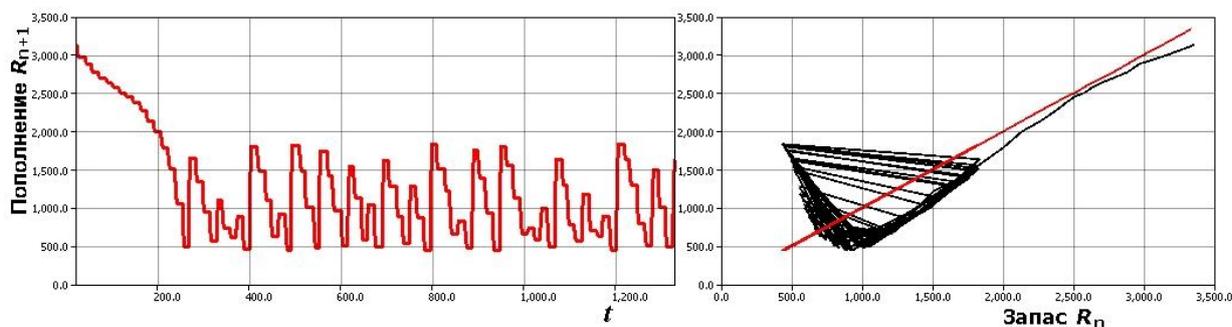


Рис. 5. Притяжение к интервальному аттрактору: временная и фазовая диаграммы

Заключение

Разработанная модель формирования пополнения позволяет описывать зафиксированные в ряде случаев [13] для промысловых популяций горбуши тихоокеанского побережья Канады изменения, связанные с существованием двух уровней численности популяции: низкого и высокого. При низкой численности популяция испытывает резкие флуктуации, но существует перспектива восстановления высокой численности при прекращении промысла. В режиме переходного хаоса невозможно предсказать, к какому из альтернативных аттракторов в результате устремится траектория, и данное свойство определяется как неопределенность относительно асимптотического состояния динамической системы.

При исследовании модели установлена возможность трансформации двух различных хаотических режимов при касательной бифуркации: переходного и устойчивого, связанного с интервальным аттрактором. Трансформация интерпретируется как следствие усиления промыслового давления и влечет длительное пребывание популяции в неблагоприятном состоянии.

Основная идея практического применения разработанных систем непрерывно-дискретных уравнений состоит в организации набора вычислительных модельных сценариев для анализа эффективности эксплуатации водных биоресурсов с учетом представлений теории этапности развития рыб [14].

Условием применения подхода является представление «стратегии» природопользования, вырабатываемой экспертами согласно некоторым внутренним правилам, применяемым для достижения приоритетной цели. Формирование сценариев на основе моделей теории пополнения запасов даст возможность рассматривать не просто динамику отдельной популяции,

но оценить концептуальные стратегии управления с точки зрения возрастания экологических рисков и регулярно отмечаемых явлений «коллапсов» рыбных запасов.

Часто подобный коллапс запасов одного вида запускает «эффект домино» по пищевым цепям экосистемы и активирует механизмы обратной связи, из-за которых численность популяции не сможет восстановиться, так как экологическая ниша занимает быстро размножающимся видом-вселенцем. Например, мелким планктоноядным рыбам тюлькам и шпротам в Черном и Каспийском морях приходится испытывать влияние размножившихся медуз.

Своевременное снижение доли изъятия и осторожный подход к организации эксплуатации биоресурсов оказываются экономически эффективнее стратегии максимизации вылова биоресурсов. Несравнимо большие потери приносит всей экономике региона вынужденно вводимый мораторий на промысел из-за критического сокращения способности к самовосстановлению истребляемых популяций.

В дальнейших исследованиях планируется применить разработанный подход для модельной оценки эффективности искусственного воспроизводства тех популяций рыб, для которых доля заводского выпуска составляет не менее 40 % ежегодного пополнения. Массовый выпуск молоди не всегда ведет к ожидаемому увеличению уловов, и определение прогнозируемого коэффициента ее промыслового возврата представляет существенную сложность. Выпуски партий молоди осетровых рыб удобно представлять в разрабатываемом подходе как предикативно выделяемые дискретные события, меняющие характер моделируемого процесса при переходе всего получающегося поколения к морскому периоду жизни.

Исследования выполнены в рамках проекта РФФИ № 14-07-00066.

Список литературы

1. Вул, Е.Б. Универсальность Фейгенбаума и термодинамический формализм / Е.Б. Вул, Я.Г. Синай, К.М. Ханин // *Успехи математических наук.* – 1984. – Т. 39, вып. 3. – С. 3–37.
2. Фейгенбаум, М. Универсальность в поведении нелинейных систем / М. Фейгенбаум // *Успехи физических наук.* – 1983. – Т. 141, вып. 2. – С. 343–374.
3. Переварюха, А.Ю. Интерпретация поведения моделей динамики биоресурсов и моментальная хаотизация в новой модели / А.Ю. Переварюха // *Нелинейный мир.* – 2012. – № 4. – С. 255–261.
4. Singer, D. Stable orbits and bifurcations of the maps on the interval / D. Singer // *SIAM journal of applied math.* – 1978. – Vol. 35. – P. 260–268.
5. Vellekoop, M. On intervals, transitivity = chaos / M. Vellekoop, R. Berglund // *The American Mathematical Monthly.* – 1994. – Vol. 101, № 4. – P. 353–355.
6. Allee, W.C. Studies in animal aggregations: mass protection against colloidal silver among goldfishes / W.C. Allee, E Bowen // *Journal of Experimental Zoology.* – 1932. – № 2. – P. 185–207.
7. Dynamical complexities in the Leslie-Gower predator-prey model as consequences of the Allee effect on prey / E. Gonzalez-Olivares [et al.] // *Appl. Math. Modell.* – 2011. – Vol. 35. – P. 366–381.
8. Yongli, C. Spatiotemporal complexity of a Leslie-Gower predator-prey model with the weak Allee effect / C. Yongli, Zh. Caidi, W. Weiming // *Journal of Applied Mathematics.* – 2013. – Vol. 2013, Article ID 535746.
9. Ricker, W. Stock and recruitment / W. Ricker // *Journal Fisheries research board of Canada.* – 1954. – Vol. 11, № 5. – P. 559–623.
10. Paar, V. Sensitive dependence of lifetimes of chaotic transient on numerical accuracy for a model with dry friction and frequency dependent driving amplitude / V. Paar, N. Pavin // *Modern Physics Letters B.* – 1996. – Vol. 10, № 4 & 5. – P. 153–159.
11. Grebogi, C. Chaotic attractors in crisis / C. Grebogi, E. Ott, J.A. Yorke // *Physical Review Letters.* – 1982. – Vol. 48, № 22. – P. 1507–1510.
12. Grebogi, C. Chaos, strange attractors and fractal basin boundaries in nonlinear dynamics / C. Grebogi, E. Ott, J.A. Yorke // *Science.* – 1987. – Vol. 238, № 4827. – P. 632–638.
13. Minto, C. Blanchard Survival variability and population density in fish populations / C. Minto, R.A. Myers // *Nature.* – 2008. – Vol. 452. – P. 344–348.

14. Еремеева, Е.Ф. Теория этапности развития и её значение в рыбоводстве / Е.Ф. Еремеева, А.И. Смирнов // Теоретические основы рыбоводства. – М. : Наука, 1965. – С. 129–138.

Поступила 24.12.2013

*Санкт-Петербургский институт
информатики и автоматизации РАН
e-mail: temp_elf@mail.ru*

A.Y. Perevaryukha

**VARIETIES OF APERIODIC DYNAMICS
IN THE EVENT-DRIVEN POPULATION MODELS**

The paper proposes a population model with the event-step structure, which includes continuous and discrete components. The dynamics of a hybrid system is analyzed in a computing environment based on the numerical solution of the sequence of Cauchy problems for the system of differential equations of generations decrease. We examine the dynamics of the functional iteration, which has two local extrema and characterizes the impermanence of the fish reproduction effectiveness. A transitional aperiodic regime is established with the possibility of attracting the trajectory to two attractors. After the bifurcation of disappearance of two nontrivial stationary points, an interval attractor arises for which a boundary crisis is possible.

УДК 519.8

А.М. Бондоловский

ОБЗОР МОДЕЛЕЙ УПРАВЛЕНИЯ ДОХОДНОСТЬЮ В ГОСТИНИЧНОМ БИЗНЕСЕ

Приводится краткое описание теории управления доходностью в гостиничном бизнесе, раскрываются основные понятия. Предлагается новая классификация процессов управления доходностью, дается обзор литературы по динамическому ценообразованию, методам прогнозирования и оптимизационным моделям, применяемым в управлении доходностью в гостиничном бизнесе. Указываются перспективные направления будущих исследований.

Введение

Приемы управления доходностью и ценообразованием возникли в индустрии пассажирских авиаперевозок в конце 70-х гг. XX в. Их достоинства в полной мере проявились при использовании в американских авиалиниях (American Airlines) в 1985 г., когда в результате одного года работы системы управления доходностью произошло увеличение дохода более чем на 14 %, а прибыль возросла на 48 % [1]. В 1990-х гг. гостиничный бизнес начал перенимать у пассажирских авиакомпаний опыт использования моделей управления доходностью и подстраивать их под свою специфику. Внедрение моделей оказалось возможным потому, что гостиничный бизнес, как и бизнес пассажирских авиаперевозок, обладает подходящим набором характеристик [2], суть которых заключается в том, что компания, желающая применять управление доходностью, должна иметь: ограниченные ресурсы (например, номера, пассажирские места, сдаваемые в аренду автомобили); продукцию либо услуги с ограниченным периодом продаж и ценностью, теряемой после определенного срока; возможность принимать заказы на будущее; низкие переменные и высокие постоянные издержки; колеблющийся спрос на свою продукцию либо услуги; возможность сегментировать рынок либо клиентов. Деятельность многих организаций обладает такими характеристиками. Поэтому в последнее время компании, которые предлагают в аренду конференц-центры, поля для гольфа, автомобили, рестораны и торговые центры, а также продают путешествия на круизных лайнерах, все больше применяют управление доходностью в своей деятельности.

Определения управления доходностью даны многими исследователями. Приведем некоторые из них на примере гостиничного бизнеса. Эль Хадад, Ропер и Джонс [3] определяют управление доходностью как важный инструмент, который соотносит спрос с предложением и максимизирует доход гостиницы путем разделения ее клиентов по разным категориям на основе их намерений по заселению и текущей емкости гостиницы. Кимез и Виртз [4] определяют этот термин как использование информационных систем и ценовых стратегий, которые в нужное время и в нужном месте сводят соответствующее количество заказов со свободным количеством соответствующих номеров. Джонси, Мичел и Сламмет [5] рассматривают управление доходностью как интегрированный, непрерывный, систематический подход по максимизации дохода от продажи номеров по изменяющимся под спрогнозированный спрос ценам. Донаги, Мак-Маан и Мак-Дауэл [6] придерживаются приблизительно того же мнения, но также выделяют важность сегментации рынка. Они определяют управление доходностью как прием максимизации дохода, который увеличивает чистый доход гостиницы через соотнесение заранее спрогнозированных свободных номеров с предопределенными сегментами рынка по оптимальной цене. Джонс и Гамильтон [7] утверждают, что управление доходностью способствует максимизации цены на номер, когда спрос превышает предложение, и максимизации загрузки, когда предложение превышает спрос, не опускаясь в цене ниже среднего уровня издержек. Все определения отмечают способность увеличения доходов компании без непосредственного регулирования издержек. Проще говоря, управление доходностью – это инструмент, который используется для увеличения доходов гостиницы. Следует обратить внимание на то, что в англоязычной литературе термину «управление доходностью» соответствует термин Revenue

Management, а в более ранние периоды – Yield Management, что может быть переведено как управление выработкой. В научной литературе произошел переход от последнего, более узкого термина, к первому [2].

Со времени начального практического успеха от использования моделей управления доходностью ученые провели достаточно много исследований на эту тему, а также обобщили и систематизировали результаты [2, 8–11]. На данный момент теоретические знания, практический опыт и программные приложения очень хорошо развиты в системах управления доходностью для авиакомпаний [12]. Меньше внимания ученые в этой области уделили гостиничному бизнесу и сфере гостеприимства. Исследования для гостиничного бизнеса фрагментарны. Имеется разрыв между теорией управления доходностью и практикой ее использования в гостиницах. Среди существующих обзоров литературы по управлению доходностью в гостиничном бизнесе следует отметить общие систематизирующие работы [2, 7, 9, 13], а также работы по прогнозной составляющей [14–16] и оптимизационной составляющей [17, 18].

Кахнеман, Кнетч и Тайлер [19, 20], Кимез [21], Виртз и др. [22] рассуждают, что, несмотря на явные преимущества моделей управления доходностью для компании-пользователя: прирост дохода, отсутствие риска при переходе к политике динамического ценообразования, эти модели могут вызывать у клиентов чувство «нечестной сделки» и вести к потере их доверия [21]. Почему клиенты могут оставаться недовольными? Управление доходностью – это система, которая контролирует потребительский спрос посредством инструментов динамического ценообразования и управления ресурсами (номера) для увеличения прибыльности. В результате работы системы цены за одинаковый срок пребывания в номере одного типа в одной гостинице могут изменяться от одного клиента к другому. Неудовлетворенность клиентов от осознания того, что один и тот же номер продается другому дешевле, в долгосрочном периоде приводит к сокращению доходов гостиницы. Поэтому Эль Хадад, Ропер и Джонс [3] указывают, что высокий прирост доходов от использования моделей управления доходностью не будет считаться успехом без оценки денежного эквивалента потери доверия клиентов к гостинице. Однако в 2008 г. в своем исследовании Палмер и Мак-Маан-Битти [23] пришли к выводу, что только степень индивидуальной осведомленности клиентов о правилах ценообразования может привести к потере доверия к отелю, а не изменяющиеся цены сами по себе. Они также заключили, что некоторые типы клиентов более восприимчивы к динамическому ценообразованию и с желанием принимают правила компании, в то время как другие не принимают, что и вызывает их неудовлетворенность от сделки. Молодые, хорошо образованные люди, которые часто совершают покупки, более других склонны доверять динамическому ценообразованию. При этом некоторые из них не только немедленно соглашаются на покупку, если предлагаемая цена ниже воспринимаемой ими справочной, но и становятся анализирующими рациональными покупателями и начинают «играть» с компаниями. Они учитывают возможное будущее движение цен при принятии решения о покупке. Для того чтобы модели управления доходностью учитывали поведение таких рациональных клиентов, Бесанко и Винстон [24] предлагают использовать подходы теории игр. Другие авторы [8, 10, 25] также соглашаются с логичностью использования приемов теории игр, однако результативность приемов пока слабо исследована.

Существует направление исследований, в котором оптимизация доходов производится при условии, что информация о спросе полностью известна, спрос превышает имеющиеся ресурсы и задача заключается в том, чтобы выбрать такие заявки, которые максимизируют доход. Литература по указанному направлению может быть найдена по ключевым словам combinatorial auctions и interval scheduling [26–28].

1. Система управления доходностью в гостинице

Управление доходностью в гостинице может быть представлена как система с взаимосвязанными элементами. Приведем схему такой системы, ее связь с входными и выходными данными и внешними факторами, влияющими на ее деятельность. Общая структура схемы взята из статьи Иванова и Жечева [13] и доработана автором (рисунок). Обозначения ПО и УД означают программное обеспечение и управление доходами соответственно.

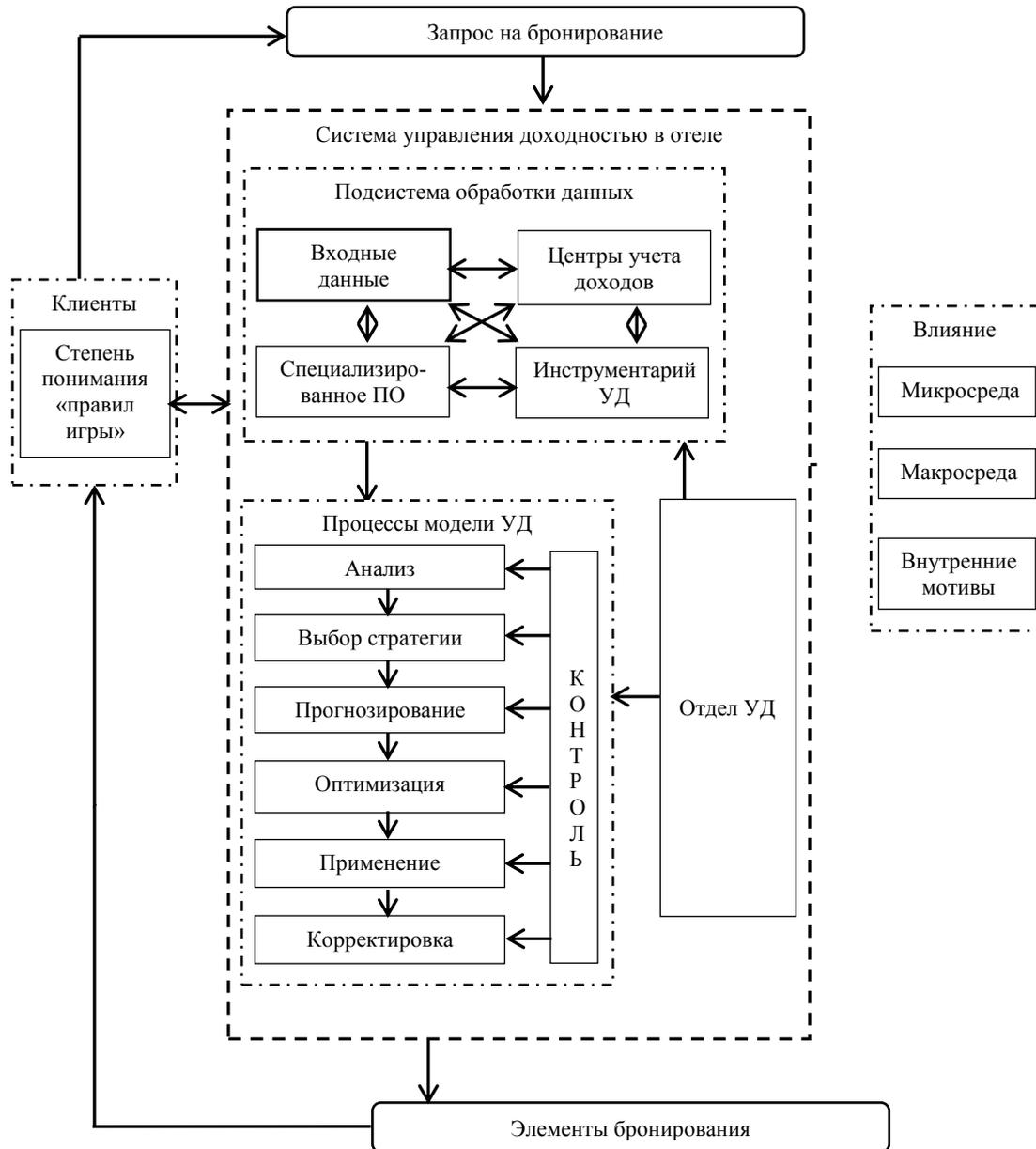


Схема системы управления доходностью гостиницы

Запрос на бронирование номера поступает от клиента и регистрируется в системе. В состав системы входит подсистема обработки данных с четырьмя тесно связанными между собой элементами: входными данными, центрами учета доходов, специализированным ПО и инструментарием УД, а также подсистема процессов модели управления доходностью и отдел УД. Входные данные несут всю информацию о заказе и возможную информацию о клиенте, а специализированное ПО регистрирует заказ и начинает его обработку с помощью определенной стратегии. Если в гостинице есть только один центр учета доходов, то он отвечает за основной доход – от продажи номеров, а если несколько таких центров, то каждый отвечает за соответствующую услугу: спа- и фитнес-залы, ресторан и бар, игорную комнату и пр. Подсистема процессов обрабатывает конкретный заказ и выдает информацию о статусе заказа, количестве и типах номеров, датах пребывания и ценах. Отдел УД – подразделение гостиницы, отвечающее за управление доходностью, прямо или косвенно утверждает результат, и он уходит к клиентам. Результат сам по себе и принцип действия системы влияют на восприятие клиентами системы ценообразования и гостиницы в целом, на их намерения осуществлять заказы в будущем.

На систему управления доходностью постоянно влияют внешние макро- и микрофакторы, а также ситуация внутри управляющей компании.

Выбор правильного инструмента, с помощью которого модель управления доходностью будет стараться максимизировать доход, очень важен. Существует достаточно много таких инструментов. В основном их можно разделить на ценовые и неценовые. К ценовым инструментам относят ценовую дискриминацию, возведение ценовых барьеров, динамическое ценообразование, гарантию наименьшей цены и другие методы, прямо влияющие на цену. Неценовые инструменты не изменяют цену непосредственно, а имеют отношение к управлению ресурсами, процессу контроля избыточного бронирования и продолжительности пребывания и к технике, гарантирующей наличие свободных номеров. Очень часто на практике применяют оба типа инструментов одновременно.

Ниже приведен краткий обзор литературы по основным инструментам управления доходами и более детальный – по динамическому ценообразованию. Начнем с неценовых инструментов. В общем виде проблему управления ресурсами гостиницы осветили Пулман и Роджерс [29]. Задачи, которые решаются в управлении ресурсами, они разделили на стратегические и краткосрочные. Стратегические задачи связаны с физическим увеличением числа номеров, возможностью изменять объемы ресурсов – количество номеров разных типов – в зависимости от спроса, нахождением оптимальных объемов. Краткосрочные задачи отвечают за каждодневную загрузку номеров, время заселения и выселения, время, затрачиваемое на обслуживание, применение к рабочим процессам моделей теории очередей и линейного программирования.

Процесс контроля избыточного бронирования основан на предположении, что по тем или иным причинам часть клиентов в гостиницу не заселится. Желая защитить себя от потерь, гостиница продает большее число номеров. По этой причине важно четко планировать избыточный уровень. Данную тему исследовали Хаджиникола и Панаи [30], Иванов [31, 32], Койд и Ишии [33], Нетесин и Шумский [34].

Меньше внимания в литературе уделено контролю продолжительности пребывания. Обычно закрепляется минимальное количество ночей пребывания. Это делается с целью защиты от краткосрочных заказов в период большого спроса и увеличения срока пребывания в период низкого спроса. Данную тему исследовали Кимез и Чейз [35] и Винод [36].

В основе ценовых инструментов лежит ценовая дискриминация, которая основывается на восприимчивости к уровню цен разных категорий клиентов, например туристов и деловых людей [4, 37, 38]. Поэтому один и тот же номер им продают по разным ценам. Чтобы избежать перехода клиентов с высоких цен на низкие, устанавливаются ценовые барьеры [39]. Они задаются определенными условиями, которые существуют при продаже того или иного номера. Например, гостиница может продавать номера только на определенные дни недели и/или на определенную продолжительность пребывания, вести строгую политику отмен заказов или продавать номера только определенным типам клиентов.

Иногда гостиницы гарантируют клиентам цену, наименьшую из существующих на рынке. Это значит, что если клиент в течение 24 ч в другой гостинице найдет похожий номер по меньшей цене, то они сравняют цены. Этот подход исследовался Карвелом и Куаном [40], Демирсифтки и др. [41].

Динамическое ценообразование – наиболее распространенный и развитый интеллектуальный ценовой инструмент [23]. Посредством него гостиница предлагает цены, которые соответствуют текущему уровню спроса и загрузки и реагируют на их изменения. Динамическое ценообразование отличается от оптимизационных моделей управления ресурсами, которые регулируют доступность номеров, тем, что воспринимает цену как переменную в функциональной зависимости спроса от цены [10]. Цена – одна из наиболее эффективных переменных, изменяя которую, менеджеры могут поощрять или ограничивать спрос в краткосрочном периоде, регулировать запасы. На цену, в свою очередь, влияют момент совершения заказа, изначальный уровень спроса, наличие номеров и другие факторы.

Быстрое развитие информационных технологий, рост электронной коммерции и всеобщее использование Интернета привели к тому, что в первом десятилетии XXI в. инструменты динамического ценообразования стали активным компонентом систем управления доходностью [42–44]. Основные причины расширения использования моделей следующие: 1) цифровой

формат обработки данных очень легко позволяет собирать ценную информацию о спросе, количестве ресурсов, ценах конкурентов и обрабатывать ее в реальном времени; 2) практически исчезли издержки на перепечатывание ценников и информирование клиентов об изменении цен [45]; 3) клиентам очень легко следить за изменениями цен и заниматься поиском. Кроме того, когда специалисты пришли к пониманию, что такие классические задачи исследования операций, как оптимизация ресурсов и управление запасами или управление потоками в сетях массового обслуживания, не могут быть отделены от ценовых решений, модели динамического ценообразования стали включаться в системы управления доходностью еще чаще и сейчас являются наиболее важной их составляющей.

Некоторые ученые, например Бойд и Билеган [46], склонны отделять модели динамического ценообразования от моделей управления доходностью. Однако они все же признают их взаимосвязь и схожесть в определенных случаях. Более того, они считают, что в случае, когда оказывается только один тип услуг, имеется один тип номера в гостинице, обе модели равнозначны.

Возможность обработать изначально прогнозируемый спрос как инструментами динамического ценообразования, так и оптимизационными моделями управления доходностью привела к тому, что названия методов стали взаимозаменяемыми [46]. Ван Райзин и Галлего [47] указывают на естественное родство между ценовыми и оптимизационными моделями. Если цена рассматривается как переменная, которую можно постоянно контролировать, то принимать решение об отказе заказа не нужно. Стоит только поднять цену достаточно высоко. Проблемы управления доходностью сквозь призму динамического ценообразования рассматривали также Ладани и Арбел [48], Галлего и Ван Райзин [44, 47], Фенг и Галлего [43], Ю [49].

Модели динамического ценообразования очень хорошо исследованы в сфере пассажирских авиаперевозок. Поэтому ученые, когда делают обзоры литературы по динамическому ценообразованию, чаще всего используют результаты для авиакомпаний. Схожесть условий продажи номеров отеля и мест в салоне самолета позволяет им говорить о равной применимости моделей в двух отраслях и описывать лишь условия перехода применения модели из одной сферы в другую. Везефорд и Бодили [11], Мак-Джил и Ван Райзин [12] в общих обзорах моделей управления доходностью отдельно остановились и на моделях динамического ценообразования. Мак-Джил и Ван Райзин, например, особо выделили работы [50 – 52], которые объединяют задачи поиска цен и управления ресурсами. Гайман [50] предпринимает попытку объединить вопросы цены и емкости. Везефорд [52] рассматривает среднюю величину нормально распределенного спроса как линейную функцию от цены.

Ряд исследователей рассматривают только модели динамического ценообразования и предлагают их классификацию. Битран и Калдентей [8] формулируют общую задачу максимизации дохода компании, которая владеет ограниченным и теряющим свою ценность набором ресурсов и работает с чувствительными к цене клиентами. К этой задаче они применяют различные модели динамического ценообразования, разделив их на детерминированные и стохастические. В каждой из двух категорий моделей они рассматривают случаи одного и нескольких типов продуктов и предлагают оптимальные решения двух видов: с одной статичной ценой на весь горизонт планирования и с динамически изменяющимися ценами. Элмаграби и Кескиноак [10] разделяют модели динамического ценообразования по категориям в зависимости от следующего: возобновляемые или невозобновляемые ресурсы, зависимый или независимый спрос, недальновидные или рациональные потребители. Они пришли к выводу, что все ценовые и оптимизационные модели можно разделить на две категории: с невозобновляемыми ресурсами, независимым спросом и недальновидными или рациональными потребителями; с возобновляемыми ресурсами, независимым спросом и недальновидными потребителями.

В заключение отметим, что при поиске той или иной оптимальной ценовой стратегии в модели часто включают ценовые ограничения. Среди наиболее общих выделяют:

- выбор цены из заданного множества [53–56];
- ограничение сверху на число различных значений цены [43];
- заранее определенные стратегии изменения цен: снижение или повышение цены во времени, специальные предложения по определенным дням и т. п. [42];
- ценовые ограничения на комплекс продуктов;
- установление цен на основе издержек.

2. Процессы модели управления доходностью

Исследователи выделяют разное количество процессов в моделях управления доходностью. Например, Грантер и др. [57] описывают модели восемью процессами: осведомленность о клиентах, сегментация рынка, внутренний анализ, конкурентный анализ, прогнозирование спроса, анализ каналов распределения, динамическое ценообразование и управление запасами и каналами распределения. Эмексиз и др. [58] предлагают пять процессов для описания модели: подготовка, анализ спроса и предложения, применение моделей управления доходностью, оценка деятельности системы управления доходностью, наблюдение и внесение изменений в модель. В свою очередь, основываясь на анализе литературы и опыте работы в гостиничном бизнесе, автор считает, что шесть процессов: анализ, выбор стратегии, прогнозирование, оптимизация, контроль, корректировка – способны в полной мере описать работу модели управления доходностью.

Компонент «анализ» включает в себя обработку входных данных и всей сопутствующей информации, а также анализирует спрос. На этапе выбора стратегии решается, какие методы прогнозирования и оптимизации ресурсов будут использованы. «Прогнозирование» и «оптимизация» – два наиболее существенных и необходимых компонента всей модели [59]. При переходе от прогнозирования к оптимизации происходит соотнесение вероятного спроса с набором ограниченных ресурсов. Поэтому важно, чтобы ошибка прогноза была наименьшей и вносила минимальную погрешность в результат работы модели. Отдел управления доходностью отвечает за компоненты «контроль» и «корректировка», следит за достижением основной цели – максимизации дохода. Ошибки и недочеты, выявленные в процессе контроля, исправляются и не проявляются при дальнейшем использовании модели.

Ниже подробно опишем компоненты «прогнозирование» и «оптимизация».

2.1. Прогнозирование

Прогнозирование – существенная и необходимая часть модели управления доходностью. Ее задача – определить будущий спрос на номера. Качество модели существенно зависит от точности прогноза. Без точного прогноза информация о цене и доступности номеров, выдаваемая моделью управления доходностью, может оказаться весьма ошибочной. Полт [60] подсчитал, что сокращение ошибки прогноза на 20 % ведет к однопроцентному увеличению дохода, получаемого от использования модели управления доходностью. Перед каждой задачей прогнозирования исследователю нужно ответить на следующие вопросы: что прогнозировать, какую степень дезагрегации объекта прогнозирования выбрать, ограничивать удовлетворение спроса или не ограничивать, какой исторический период для сбора входных данных использовать, на какой интервал времени в будущем распространяется прогноз, какой метод прогнозирования выбрать, какую точность прогноза считать приемлемой.

Основным объектом прогнозирования в отельном бизнесе является спрос: количество заказов на номера и отмен совершенных заказов, продолжительность заселения. Природа бронирования номеров такова, что заказы обычно приходят за дни, недели или месяцы до заселения. По этой причине прогнозируемому спросу свойственны две временные характеристики: время совершения заказа и время заезда в гостиницу. Каждый заказ также характеризуется типом номера и продолжительностью заселения. Природа отмен совершенных заказов подобна, за исключением двух важных особенностей. Отменить можно только совершенные заказы. Их отменяют обычно за некоторое число дней до заселения. Величина, показывающая, за сколько дней до заселения был отменен заказ, необходима при прогнозе или имитации отмен. Разница между совершенными заказами и отменами называется чистыми заказами. Перед выбором прогнозной модели также решают, прогнозировать спрос только в пределах существующих ресурсов либо учитывать весь спрос, не ограниченный емкостью отеля или ценами.

Прогнозировать можно данные разной степени агрегации: совокупные, частично агрегированные и полностью дезагрегированные данные. Выбор степени агрегации зависит от типа имеющихся данных и метода прогноза. При совокупном подходе сначала делается прогноз по всей гостинице, а затем посредством исторических вероятностей делится по категориям. При полностью дезагрегированном подходе прогноз делается по каждой категории, а затем при

необходимости объединяется. В своей публикации Везефорд, Кимез и Скот [61] показали, что полностью дезагрегированный прогноз обычно дает лучшие результаты, чем частично агрегированный или совокупный прогноз.

Известно, что для спроса в гостиничном бизнесе характерна высокая степень сезонности. Если использовать малый период исторических данных, например 8–12 недель, то сезонность может быть не учтена, а если большой, то сезонность может быть лучше передана, но тогда нужно подобрать точный период исторических данных. Следует помнить, что относительно большой период исторических данных может сделать прогноз недостаточно достоверным. Период, на который строится прогноз, называется периодом упреждения. Выделяют долгосрочный и краткосрочный периоды упреждения. В долгосрочном периоде конечный спрос прогнозируется значительно ранее дня заезда (до года). Краткосрочный период прогноза изменчив, но обычно не превышает трех месяцев. На практике большинство заказов совершается на протяжении 60 дней до дня заезда.

В своей диссертации Ли [62] выделяет три типа методов прогнозирования: исторические, прогрессивные и объединенные. К историческим относят методы экспоненциального сглаживания, скользящего среднего, использования спроса такого же дня в прошлом году, авторегрессии и методы Бокса – Дженкинса ARMA и АРИМА. Метод ARMA объединяет авторегрессионный метод и метод скользящего среднего и применим только к стационарным временным рядам. Методы АРИМА являются расширением методов ARMA для нестационарных временных рядов. Для получения конечного прогноза исторические методы используют только данные прошлых периодов, общее число заездов по конкретному предыдущему дню за определенный период в прошлом. Интересно, что в ранних исследованиях чаще использовали простые методы, тогда как в поздних – более сложные. Макридакис и др. [63] подытожили, что такие статистически усложненные методы, как АРИМА, в большинстве случаев не превосходят более простые.

Прогрессивные методы на момент прогноза учитывают уже совершенные заказы на будущие дни. Эти методы называют еще методами «подхвата», потому что ожидаемое число заказов, которое будет «подхвачено» в отрезке времени между рассматриваемым днем и днем заселения, прибавляется к текущему, уже имеющемуся числу заказов. В прогрессивных методах выделяют добавочный и мультипликативный методы «подхвата» заказов. Согласно Везефорд и Кимез [64] в добавочном методе количество уже существующих заказов не зависит от конечного числа заездов, в то время как в мультипликативном методе число будущих заказов зависит от уже существующих.

Объединенные методы используют наилучшие свойства исторических и прогрессивных методов и объединяют их либо путем взвешенного усреднения, либо регрессионными методами. Метод нейронных сетей также относят в эту группу. Такие ученые, как Филдс и Орд [65], Бен-Акива [66], считают, что объединенные методы дают наиболее точные результаты прогноза. Не будем подробно останавливаться на методах прогноза, а перенаправим читателя к результатам исследований других авторов (табл. 1).

Таблица 1

Методы прогноза в литературе

Исторические	Экспоненциальное сглаживание	Бюргер и др. [14], Чен и Качани [15], Раджопадх и др. [67], Везефорд и Кимез [64], Юксель [68], Пхумчасри и Монгколкул [16]
	Скользящее среднее	Бюргер и др. [14], Везефорд и Кимез [64], Юксель [68]
	Авторегрессионные методы	Бюргер и др. [14], Лим и Чан [69], Лим, Чанг и МакАлер [70], Юксель [68]
Прогрессивные	Добавочный метод	Чен и Качани [15], Везефорд и Кимез [64]
	Мультипликативный	Везефорд и Кимез [64]
Объединенные	Регрессионные методы	Бюргер и др. [14], Чен и Качани [15], Везефорд и Кимез [64]
	Взвешенное усреднение	Чен и Качани [15]

Правильный выбор прогнозной модели очень важен. Чаще всего ее точность является главным критерием при выборе модели. Существует несколько инструментов оценки точности прогноза. Оценка по среднему абсолютному отклонению – наиболее простой и применимый способ. По нему усредняют абсолютные значения прогнозных ошибок, а затем выбирают вариант с его минимальным значением либо стараются минимизировать. Используют также средний процент ошибки, среднее абсолютное процентное отклонение, среднее квадратичное отклонение и другие оценки [16]. Достаточно полное сравнение методов оценок прогноза провели Армстронг и Колопи [71].

Ученые по-разному оценивают эффективность прогнозных методов. Везефорд и Кимез [64], используя реальные исторические данные сетей гостиниц Чойз Хотелс и Мариот Хотелс и сравнив эффективность прогнозных методов, пришли к заключению, что экспоненциальное сглаживание, скользящее среднее и метод подбора уже совершенных заказов дают наиболее точные прогнозы. В свою очередь, Филдс и Орд [65] провели эксперименты, результат которых показал, что объединенные модели дают наибольшую точность. Захари, Гаяр и Атя [72] в своих экспериментах обнаружили, что добавочный метод подбора уже существующих заказов дает более точные результаты, чем мультипликативный метод. Скинаарс [73] заключил, что при высоко изменчивых входных данных метод переноса спроса из такого же дня в прошлом превосходит другие популярные методы. Несмотря на некоторую разницу в оценках, все ученые сходятся во мнении, что разные модели должны применяться к разным типам данных, отличающимся по сезону, группе клиентов и другим параметрам.

Для получения наиболее точных результатов прогноза некоторые исследователи предлагают внедрять опыт и знания экспертов в прогнозные модели, объединяя их с математической частью. Это направление исследований весьма актуально в наши дни. Общие наблюдения показывают, что менеджеры гостиницы в состоянии дать очень точный прогноз на две-три недели вперед [67]. Человеческая оценка особенно полезна при учете влияния внешних событий и мероприятий на будущий спрос.

2.2. Оптимизация

Оптимизационная часть модели управления доходностью чрезвычайно важна. Именно в ней решается задача максимизации дохода гостиницы и определяются оптимальные цены и количество номеров к продаже. С учетом разных типов номеров, ценовых категорий и длительностей заездов задача становится сложной. Поэтому важно на входе в оптимизационную часть иметь прогноз высокой точности. Он помогает получить более достоверные результаты оптимизации и соответственно большее увеличение дохода. Все оптимизационные модели изначально развивались для сферы пассажирских авиаперевозок, затем из-за схожести математических моделей и сферы применения перешли в гостиничный бизнес. Поэтому часто в обзорах литературы исследователи описывают результаты по моделям в авиаперевозках и используют соответствующую авиаперевозкам терминологию. Далее мы также будем ею пользоваться.

Модели управления посадочными местами составляют ядро оптимизационных моделей в авиаперевозках [9]. С их помощью до времени отлета решается задача соотношения ограниченного числа посадочных мест с запросами клиентов, готовых платить разные цены. Цены каждого класса каждого отрезка пути подготовлены заранее. Цель задачи – найти оптимальные цены и количество мест к продаже в разных классах, которые ведут к максимизации дохода. Последующее принятие или отклонение каждого поступающего заказа основывается на полученном решении задачи. Очевидно, что в определенный момент времени более выгодно отказаться от запроса на место в низком ценовом классе в надежде продать это же место в более высоком ценовом классе. Основными методами управления посадочными местами являются: метод «отрезок пути – класс» (его еще называют «управление единичным ресурсом»), который каждый отрезок пути оптимизирует отдельно, и метод «пункт отправления – пункт прибытия», родственной первому, но оптимизирующий всю сеть отрезков пути в целом. Под отрезком пути понимается прямой перелет между двумя точками без остановки, а под маршрутом – любой возможный набор смежных отрезков пути. Оба метода могут иметь статистические и динамические решения. Статистические решения определяют оптимальные цены и количество мест один раз до старта продаж, используя прогноз спроса, полученный в этот момент времени.

Подразумевается, что к моменту прогноза имеется достаточно полная информация о спросе. Задача определения цен и количества мест может решаться периодически в моменты поступления обновленной информации о заказах и наличии свободных ресурсов. При динамическом подходе постоянно контролируется состояние процесса поступления заказов и каждый раз при поступлении запроса принимается решение о принятии либо отклонении заказа. Не будем подробно описывать методы и соответствующие решения. Достаточно подробно их описали Везефорд [74], Мак-Джил и Ван Райзин [12], Бойд и Билеган [47], Пак и Пирсма [75]. В данном разделе обсудим лишь основные понятия методов и проведем аналогии с отельным бизнесом.

Используя метод «отрезок пути – класс», авиалинии создают набор услуг, известных как классы. Например, авиакомпания продает четыре класса: А, Б, В и Г. Класс А имеет самые высокие цены и никаких ограничений на покупку и возврат билетов. В классе Г цены низкие, но билеты закреплены за определенными днями и деньги на них не возвращаются. На классы Б и В устанавливаются умеренные цены и ограничения. На каждый отрезок пути в каждом классе модель управления доходностью предлагает к продаже определенное количество билетов, и пассажир может забронировать билеты класса Б на отрезок пути 1, соединенный с отрезком пути 2, только если билеты класса Б будут доступны на оба отрезка. Рассмотрим ситуацию с двумя отрезками пути: ТЧК1–ТЧК2 и ТЧК2–ТЧК3, каждый из которых имеет только одно свободное место. Только два пассажира желают совершить перелет. Один готов заплатить \$70 за класс А на отрезок пути ТЧК1–ТЧК2, а другой готов заплатить \$210 за класс А при перелете по двум отрезкам пути ТЧК1–ТЧК2 и ТЧК2–ТЧК3. В методе «отрезок пути – класс» места доступны к продаже только тогда, когда доступны и отрезок пути и класс одновременно. Становится невозможным заблокировать продажу класса А за \$70, пока еще открыта возможность покупки класса А за \$210. Отказывая пассажирам, подобным второму, авиалинии теряют возможность получать доход от объединения двух отрезков пути.

Литлвуд [76] первым предложил статистические решения с двумя классами в методе «отрезок пути – класс». Он предлагает закрывать класс низкой ценовой категории в случае, когда доход от продажи очередного места в этом классе превзойден ожидаемым доходом от продажи того же места, но в более высоком классе. Белобаба [77] дорабатывает подход Литлвуда и предлагает гнездовой подход на несколько классов. Подход получил название «ожидаемый предельный доход с посадочного места» или ОПДМа. Гнездовой подход для решения нескольких классов независимо представили Карри [78] и Волмер [79]. Карри считает, что распределение спроса непрерывное, а Волмер – дискретное. Подход Брумеля и Мак-Джила [80] получил название ОПДМб и допускает как непрерывное, так и дискретное распределение спроса.

К динамическим решениям метода «отрезок пути – класс» относят модель динамического программирования Ли и Херша [81], в которой спрос на каждый класс моделируется неоднородным пуассоновским процессом и решение принимается в предположении, что процесс является марковским. Кливегт и Папаставро [82] показали, что проблема может быть сформулирована как динамическая стохастическая задача о ранце. Субраманиан, Стилдхэм и Лаутенбахер [83] добавили в модель Ли и Херша учет отмен заказов.

Мультистадийная задача стохастического программирования максимизации доходов авиакомпании по одному отрезку пути с большим числом классов была предложена Вилиамсом [84, 85]. Задача трансформируется под специфику управления доходами гостиницы. Под отрезком пути понимается вся гостиница, под местами разных классов – номера разных типов, под датой вылета – дата заселения на период времени. Задача Вилиамса [84] может быть применена для нахождения оптимальных цен и количеств номеров каждого типа для каждого периода времени в горизонте планирования. Подробно сформулируем задачу.

Горизонт бронирования делится на T периодов времени. В каждом периоде времени $t = 1, \dots, T - 1$ номера бронируются на период времени T . Номерной фонд гостиницы состоит из номеров трех типов $i = 1, 2, 3$. Типы 1 и 2 и 2 и 3 смежные. В период времени T выставляются к продаже n_i номеров типа i и r_i процентов номеров этого типа может быть трансформировано в номера смежных типов. Цена на номер типа i на период времени T платится в период времени t , $0 \leq t \leq T - 1$, и может принимать одно из следующих значений: $c_{t,i,1}, \dots, c_{t,i,O_t}$, где O_t есть количество ценовых опций в период времени t .

Значение спроса на период времени T – это количество номеров каждого типа. Предполагается, что спрос не определен и что его значения зависят от цены. Допустим, что с помощью прогноза в период времени t на период времени T получено S_t сценариев спроса. В то время как значения спроса зависят от цены, предполагается, что сценарии спроса от нее не зависят.

Считается, что в период времени t сценарии спроса являются независимыми и формируют полную систему событий. Пусть вероятность сценария s в период времени t , $1 \leq s \leq S_t$, равна $p_{t,s}$ и $\sum_{s=1}^{S_t} p_{t,s} = 1$.

Задача подразумевает построение дерева сценариев. Дерево строится из $T + 1$ уровней. Уровни обозначаются $t = 0, 1, \dots, T$, и каждый из них состоит из числа узлов. Каждый узел (t, s) уровня t связан со сценарием спроса s в период времени t , $t = 0, 1, \dots, T$, $s = 1, \dots, S_t$. Уровень 0 состоит из искусственного узла $(0, 0)$, где 0 – это искусственный сценарий, который осуществляется с вероятностью 1 в период времени 0. Предполагается, что для каждого узла $(t + 1, b)$ существует только одна дуга $((t, a), (t + 1, b))$, которая означает, что сценарий b в период времени $t + 1$ происходит после сценария a в период времени t , $t = 0, 1, \dots, T - 1$. Это предположение делает отношения предшествования между узлами древовидными. Если существует дуга $((t, a), (t + 1, b))$, то узел (t, a) называется предком узла $(t + 1, b)$.

Каждому узлу (t, s_t) уровня t соответствует уникальная последовательность сценариев $v = ((0, 0), (1, s_1), (1, s_2), \dots, (t, s_t))$, которая заканчивается в этом узле, $s_\tau \in \{1, \dots, S_\tau\}$, $\tau = 1, \dots, t$. Если осуществился переход к периоду времени t , то сценарий спроса этого периода реализовался и он известен. При нахождении в период времени 0 вероятность того, что последовательность сценариев $v = ((0, 0), (1, s_1), (1, s_2), \dots, (t, s_t))$ приведет к сценарию спроса s_t в период времени t , определяется как $P_v = \prod_{\tau=1}^t p_{\tau, s_\tau}$. Пусть V_t обозначает множество всех последовательностей сценариев, которые заканчиваются в узлах уровня t , $t = 0, 1, \dots, T$. Согласно древовидным отношениям предшествования $|V_t| = S_t$.

Предположим, что для каждой последовательности сценариев $v \in V_t$ в период времени t спрос на номера типа i по цене o на период времени T известен или спрогнозирован и обозначается $d_{v,i,o}$.

Опишем следующие переменные:

$x_{v,i,o}$ – количество номеров типа i для продажи в период времени t на период времени T по цене o при условии, что последовательность сценариев $v \in V_t$ реализовалась, $0 \leq t \leq T - 1$;

$y_{v,i,o}$ – вспомогательная переменная; $y_{v,i,o} = 1$, если $x_{v,i,o} > 0$, и $y_{v,i,o} = 0$, если $x_{v,i,o} = 0$, $v \in V_t$, $0 \leq t \leq T - 1$;

$z_{v,i}$ – вспомогательная переменная, которая отражает общее количество номеров типа i на период времени T , проданных по пути последовательности сценариев v , $v \in V_t$, $0 \leq t \leq T$.

Сформулируем детерминированную модель следующим образом:

$$\max \sum_{t=0}^{T-1} \sum_{v \in V_t} \sum_{o=1}^{O_t} P_v c_{t,i,o} x_{v,i,o} \quad (1)$$

при условии, что

$$\sum_{o=1}^{O_t} y_{v,i,o} = 1, v \in V_t; i = 1, 2, 3; t = 0, \dots, T - 1; \quad (2)$$

$$x_{v,i,o} \leq d_{v,i,o} y_{v,i,o}, v \in V_t; i = 1, 2, 3; o = 1, \dots, O_t; t = 0, \dots, T - 1; \quad (3)$$

$$z_{v,i} = \sum_{o=1}^{O_1} x_{0,i,o}, v \in V_1; i = 1, 2, 3; \quad (4)$$

$$z_{vi} = z_{parent(v),i} + \sum_{o=1}^{O_t} x_{parent(v),i,o}, v \in V_t; i = 1, 2, 3; t = 2, \dots, T; \quad (5)$$

$$z_{v,1} \leq \left(n_1 + \left\lfloor \frac{r_2 n_2}{100} \right\rfloor \right), v \in V_T; \quad (6)$$

$$z_{v,2} \leq \left(n_2 + \left\lfloor \frac{r_1 n_1 + r_3 n_3}{100} \right\rfloor \right), v \in V_T; \quad (7)$$

$$z_{v,3} \leq \left(n_3 + \left\lfloor \frac{r_2 n_2}{100} \right\rfloor \right), v \in V_T; \quad (8)$$

$$z_{v,1} + z_{v,3} \leq \left(n_1 + n_3 + \left\lfloor \frac{r_2 n_2}{100} \right\rfloor \right), v \in V_T; \quad (9)$$

$$z_{v,1} + z_{v,2} + z_{v,3} \leq n_1 + n_2 + n_3, v \in V_T; \quad (10)$$

$$x_{v,i,o} \in Z_+, v \in V_t; i = 1, 2, 3; o = 1, \dots, O_t; t = 0, \dots, T-1; \quad (11)$$

$$y_{v,i,o} \in \{0,1\}, v \in V_t; i = 1, 2, 3; o = 1, \dots, O_t; t = 0, \dots, T-1; \quad (12)$$

$$z_{v,i} \in Z_+, v \in V_t; i = 1, 2, 3; t = 0, \dots, T. \quad (13)$$

Целевая функция (1) представляет собой общий ожидаемый доход от продажи номеров в периоды времени $t = 0, 1, \dots, T-1$ на период T . Равенства (2) обеспечивают то, что в любой из периодов T для каждого типа номера выбирается только одна ценовая опция. Неравенства (3) гарантируют, что для любой последовательности сценариев и любой ценовой опций количество проданных номеров каждого из трех типов не превышает спроса на эти номера в случае, если последовательность реализовалась. Равенства (4) и (5) представляют рекурсивные вычисления значений переменных z через значения переменных x . Неравенства (6)–(10) объявляют верхние границы на общее число номеров каждого типа и их комбинации в период времени T .

Необходимость исправления слабых сторон метода «отрезок пути – класс» привела к появлению сетевого метода «пункт отправления – пункт прибытия». Метод более эффективен при работе с заказами с пересадками, так как одновременно оптимизирует всю сеть отрезков пути, предлагаемых авиакомпанией. Один из приемов этого метода – заранее определить пропорции распределения ожидаемого дохода всего пути от пункта отправления до пункта прибытия между отрезками пути и применить метод «отрезок пути – класс» к каждому отрезку. Инструменты математического программирования достаточно хорошо справляются с комбинаторными аспектами подобной задачи. Поэтому Гловер и др. [86], Талури и Ван Райзин [87] и многие другие формулируют задачу сетевых маршрутов как задачу линейного программирования следующим образом:

$$\max \sum_{i \in I} r_i x_i$$

при условии, что

$$\begin{aligned} \sum_{i \in I(l)} x_i &\leq c_l, \quad l \in L; \\ x_i &\leq d_i, \quad i \in I; \\ x_i &\geq 0, \quad i \in I, \end{aligned}$$

где I – множество всех пар «маршрут – класс»; r_i – цена одного места для пары i «маршрут – класс»; переменная x_i – количество заказов для пары i «маршрут – класс»; $I(l)$ – множество пар «маршрут – класс» для отрезка l ; L – множество отрезков в сети/маршруте; c_l – емкость отрезка l ; d_i – ожидаемое количество заказов для пары i «маршрут – класс». Задача состоит в определении оптимального количества заказов, которое приводит к максимизации дохода.

Наиболее используемые подходы в методе «пункт отправления – пункт прибытия» – это виртуальное гнездование и метод цены клиента. Подход виртуального гнездования близок методу «отрезок пути – класс», но позволяет управлять продажей мест по величине продаваемой стоимости, а не по классу. Создавая «виртуальные корзины» запасов, основанные на стоимости, а не на классе, подход устраняет главное неудобство метода «отрезок пути – класс». Рассмотрим описанный выше пример с двумя пассажирами и двумя отрезками пути. Создаются две виртуальные корзины. Для более прибыльного запроса создается корзина 1, а для менее прибыльного – корзина 2 (табл. 2). Затем каждый отрезок пути пары «маршрут – класс» закрепляется за виртуальной корзиной. Так как по одному месту свободно в каждом из отрезков пути, запасы (места) будут «посажены» в корзину для более прибыльных запросов – корзину 1. Для максимизации дохода запрос стоимостью в \$70 за класс А на отрезок пути ТЧК1–ТЧК2 будет направлен в корзину с нулевыми запасами мест – корзину 2, а запрос стоимостью \$210 за весь маршрут – соответственно в корзину, где места доступны, – корзину 1.

Таблица 2

Доступные места, посаженные по виртуальным корзинам

Корзины	Доступные места для отрезков пути	
	ТЧК1–ТЧК2	ТЧК2–ТЧК3
1	1	1
2	0	0

Подход цены клиента также похож на метод «отрезок пути – класс» и естественным образом включает в себя принцип гнездования. Для каждого отрезка пути маршрута сети устанавливается базовая цена продажи. Класс становится доступным для продажи, если общая ожидаемая плата за места класса в маршруте превышает сумму цен, установленных для каждого отрезка пути этого маршрута. В свое время подход цены клиента исследовали Вилиямсон [88], Вей [89], Талури и Ван Райзин [90].

Кроме виртуального гнездования и метода цены клиента, основанных на детерминированных инструментах математического программирования, для решения сетевой задачи выделяют еще имитационный и динамический подходы. Бертсимас и де Боер [75] впервые представили имитационный подход, который помимо детерминированной модели линейного программирования и учета стохастической природы спроса использует инструменты приближенного динамического программирования. Полноценное динамическое решение сетевой задачи впервые было получено Ченом, Гюнтером и Джонсоном [91]. Они сформулировали марковскую модель принятия решений, которая использует математическое программирование в динамической среде. Общие стохастические сетевые модели на основе марковского процесса принятия решений и нескольких типов аппроксимации предложили Ван Райзин и Талури [92]. Купер и Хомем-де-Мело [93] постарались объединить принятие решений на основе марковского процесса и подходы математического программирования.

Отметим, что формулировка задачи математического программирования в методах «отрезок пути – класс» и «пункт отправления – пункт прибытия» одинакова как для авиалиний, так и для гостиниц. Конечно, для гостиниц есть своя специфика. Например, классы мест, продаваемых авиакомпаниями, равнозначны эконом-, бизнес- или премиум-типам номеров в гостиничном бизнесе. Типы могут отличаться по физическим характеристикам номеров или по наличию дополнительных свойств, не связанных с физическими характеристиками номера: бесплатных завтраков, полотенец. Иногда разные типы могут быть установлены и на идентичные номера. Далее, если для задач авиалиний в методе «пункт отправления – пункт прибытия» рассматривалась сеть маршрутов, состоящая из отдельных рейсов, то для гостиниц рассматривается последовательность из ночей пребывания. В задаче линейного программирования L представляет собой набор пар «дата – тип номера» в фиксированном временном интервале, I – набор троек «дата заезда – продолжительность пребывания – тип номера», $I(l)$ – набор троек «дата заезда – продолжительность пребывания – тип номера» для пары l «дата – тип номера». Например, заезд 21 марта в номер бизнес-класса на три ночи будет соответствовать индексу $i \in I$, и если $l \in L$ соответствует пара «дата – тип номера» номера бизнес-класса на 22 марта, то $i \in I(l)$. Параметр c_l обозначает количество номеров пары l «дата – тип номера». Обычно он со временем не изменяется. Параметр d_i – ожидаемое количество номеров тройки i «дата заезда – продолжительность пребывания – тип номера». Переменная x_i характеризует количество заказов на номера типа i , которое в случае продажи по цене r_i должно быть принято.

Метод «отрезок пути – класс» для гостиничного бизнеса назовем «дата – тип номера». Он определяет доступность типа номера на каждый день и предлагает его к продаже, если тип номера свободен на все дни запроса клиента. Методу «дата – тип номера» присущи те же слабые стороны, что и методу «отрезок пути – класс». Он не может рассматривать заказы по длительности пребывания. Поэтому заказы на одну-две ночи «забывают» номера и не пропускают заказы с более длительным сроком пребывания, что ведет к неэффективному использованию ресурсов. У методов виртуального гнездования и цены покупателя тоже есть свои аналоги в отельном бизнесе. Метод виртуального гнездования разносит разные комбинации «дата заезда – продолжительность пребывания – тип номера» по разным корзинам для разных ночей, и тип номера продается при условии доступности в каждой соответствующей корзине на протя-

жении всего периода пребывания. Метод цены покупателя определяет цену покупателя на каждую ночь, и тип номера продается, если общая предлагаемая плата за номер превышает сумму предопределенных цен за весь период пребывания.

В недавнем обзоре моделей управления доходностью в гостиничном бизнесе Иванов и Жечев [13] предполагают, что среди оптимизационных моделей превалирует использование стохастического программирования (Голдман и др. [18], Лай и Нг [94], Лиу и др. [95], Лиу, Лай и Ванг [96]) и имитационных методов (Бэкер и Колиер [97], Раджопадх и др. [67], Захари и др. [98]). Методам детерминированного линейного программирования (Голдман и др. [18], Лиу Лай и Ванг [96]), целочисленного программирования (Бертимас и Шиода [99]), динамического программирования (Бадинели [100], Бертимас и Шиода [99]) и нечеткого целевого программирования (Падхи и Агарвал [101]) уделено меньше внимания, однако к ним наблюдается возрастающий интерес. Методы «цена покупателя» (Бэкер и Колиер [96]) и «ожидаемый предельный доход» (Иванов [32], Нетесин и Шумски [34]) мало используются в гостиничном бизнесе.

Заключение

Выделим некоторые актуальные направления будущих исследований. В литературе есть мнение, что разбиение спроса на категории по определенным параметрам приводит к повышению точности прогноза [2, 61]. Однако количество категорий, рассмотренных в конкретных случаях, невелико. Кимез [2], например, говорит о разбиении спроса или только по ценовой категории номера, или только по продолжительности пребывания, или по двум характеристикам одновременно. Автор предполагает, что разбиение спроса по большему числу параметров (до 10) позволит не только увеличить точность прогнозной модели, но и эффективно использовать оптимизационные модели по каждой из категорий. Это улучшит качество всей модели управления доходностью.

Разбиение спроса по категориям может приводить к разреженности исторических данных. Не все методы прогноза одинаково эффективны в условиях разреженности данных. Точность прогнозов значительно возрастет, если исследователи будут оценивать степень плотности исторических данных и применять разные методы прогноза для разреженных данных и для данных с достаточной степенью плотности.

При решении оптимизационных задач используются методы математического программирования, такие как линейное программирование и квадратичное, стохастическое линейное программирование и динамическое программирование. Автор считает, что, детализируя методы квадратичного программирования или более общие методы выпуклого программирования, можно разрабатывать новые модели и успешно решать задачи управления доходами. Эти методы учитывают некоторую нелинейность и в то же время являются достаточно эффективными в смысле использования вычислительных ресурсов и времени решения. Так, моделирование спроса на номера невозрастающей функцией от цены позволяет легко объединить задачи динамического ценообразования и управления ресурсами. Когда найдены конкретные значения коэффициента эластичности спроса и константы функции будущего спроса, цены являются неизвестными, а критерием служит максимизация суммарного моделируемого дохода при условиях ограниченных ресурсов и верхних и нижних границ на цены, посредством методов выпуклого программирования можно определять оптимальные ценовые решения. В методах квадратичного программирования квадраты переменных цен в целевой функции появляются в результате перемножения цены и эластичного спроса, линейно зависящего от цены. Так как в процессе моделирования одновременно с определением цен решается задача о нахождении количества занятых номеров всех типов, модель управляет еще и ресурсами гостиницы с учетом продолжительности пребывания каждого заселения.

Автор также заметил, что в литературе проблемы прогнозирования спроса на номера исследуются отдельно от оптимизационных задач управления ресурсами или ценами. При объединении этапов в целостную рабочую систему управления доходностью гостиницы всегда возникает ряд интересных практических задач. Кроме того, описание системы от этапа подготовки базы данных и ввода параметров до понятно оформленных результатов оптимизационных моделей, применяемых для выставления цен или принятия/отклонения поступающих заказов,

представляет интерес для менеджеров гостиниц и исследователей, дает комплексное понимание системы управления доходностью гостиницы.

Отметим также наиболее актуальные направления будущих исследований, указанные другими авторами. Иванов и Жечев [13] предполагают, что качество прогнозных моделей возрастет при развитии моделей прогноза, способных предопределять изменения спроса в связи с проведением различных событий и специальных мероприятий. Биллер и др. [102], Битран и Калдентей [8] утверждают, что внедрение в модели управления доходностью возможности учета стратегического поведения клиентов – еще один вопрос для дальнейшего рассмотрения. Какие цены должны быть установлены, если при покупке номера в гостинице клиент действует стратегически? Логично, что ответить на этот вопрос можно, рассмотрев его с точки зрения теории игр.

Список литературы

1. Nguyen, Y. Hotel revenue management: a necessary evil, but not sufficient for delivering profitably / Y. Nguyen // *hotelmarketing.com* [Electronic resource]. – 2013. – Mode of access : <http://goo.gl/pb2UfN>. – Date of access : 22.04.2014.
2. Kimes, S.E. Revenue management: a retrospective / S.E. Kimes // *Cornell hotel and restaurant administration quarterly*. – 2004. – Vol. 44. – P. 131–138.
3. El Haddad, R. The impact of revenue management decisions on customers attitudes and behaviours: a case study of a leading UK budget hotel chain / R. El Haddad, A. Roper, P. Jones // *EuroCHRIE 2008 Congress*. – Dubai : Emirates hotel school, 2008.
4. Kimes, S.E. Has revenue management become acceptable? Findings from an international study on the perceived fairness of rate fences / S.E. Kimes, J. Wirtz // *Journal of service research*. – 2003. – Vol. 6, № 2. – P. 125–135.
5. Jauncey, S. The meaning and management of yield in hotels / S. Jauncey, I. Mitchell, P. Slamet // *International journal of contemporary hospital management*. – 1995. – Vol. 4. – P. 23–26.
6. Donaghy, K. Yield management: an overview / K. Donaghy, U. McMahon, D. McDowell // *International journal of hospitality management*. – 1995. – Vol. 14, № 2. – P. 139–150.
7. Jones, P. Yield management: putting people in the big picture / P. Jones, D. Hamilton // *The Cornell hotel and restaurant administration quarterly*. – 1992. – Vol. 33, № 3. – P. 89–96.
8. Bitran, G. An overview of pricing models for revenue management / G. Bitran, R. Caldentey // *Manufacturing and service operations management*. – 2003. – Vol. 5, № 20. – P. 203–229.
9. Chiang, W. An overview of research on revenue management: current issues and future research / W. Chiang, J.C.H. Chen, X. Xu // *International journal of revenue management*. – 2007. – Vol. 1, № 1. – P. 97–128.
10. Elmaghraby, W. Dynamic pricing in the presence of inventory considerations: research overview, current practices, and future directions / W. Elmaghraby, P. Keskinocak // *Management science*. – 2003. – Vol. 49, № 10. – P. 1287–1305.
11. Weatherford, L.R. A taxonomy and research overview of perishable-asset revenue management: yield management, overbooking, and pricing / L.R. Weatherford, S.E. Bodily // *Operations research*. – 1992. – Vol. 40, № 5. – P. 831–843.
12. McGill, J.I. Revenue management: research overview and prospects / J.I. McGill, G.J. van Ryzin // *Transportation science*. – 1999. – Vol. 33. – P. 233–256.
13. Ivanov, S. Hotel revenue management – a critical literature review / S. Ivanov, V. Zhechev // *Tourism*. – 2012. – Vol. 60, № 2. – P. 175–197.
14. A practitioners guide to time-series methods for tourism demand forecasting – a case study of Durban, South Africa / C.J.S. Burger [et.al.] // *Tourism management*. – 2001. – Vol. 22, № 4. – P. 403–409.
15. Chen, C. Forecasting and optimisation for hotel revenue management / C. Chen, S. Kachani // *Journal of revenue and pricing management*. – 2007. – Vol. 6, № 3. – P. 163–174.
16. Phumchusri, N. Hotel room demand forecasting via observed reservation information / N. Phumchusri, P. Mongkolkul // *Proceedings of the Asia Pacific industrial engineering and management systems conference; Asia Pacific industrial engineering and management society*. – Phuket, 2012. – P. 1978–1985.

17. Bitran, G. An application of yield management to the hotel industry considering multiple day stays / G. Bitran, S. Mondschein // *Operations research*. – 1995. – Vol. 43, № 3. – P. 427–443.
18. Models and techniques for hotel revenue management using a rolling horizon / P. Goldman [et al.] // *Journal of revenue and pricing management*. – 2002. – Vol. 1, № 3. – P. 207–226.
19. Kahneman, D. Fairness and the assumptions of economics / D. Kahneman, J.L. Knetsch, R.H. Thaler // *Journal of business*. – 1986. – Vol. 59. – P. 285–300.
20. Kahneman, D. Fairness as a constraint on profit seeking: entitlements in the market / D. Kahneman, J.L. Knetsch, R.H. Thaler // *The American Economic Review*. – 1986. – Vol. 76, № 4. – P. 728–741.
21. Kimes, S.E. Perceived fairness of yield management / S.E. Kimes // *The Cornell hotel and restaurant administration*. – 1994. – Vol. 35, № 1. – P. 22–24.
22. Revenue management: resolving potential customer conflicts / J. Wirtz [et al.] // *Journal of revenue and pricing*. – 2003. – Vol. 2, № 3. – P. 216–226.
23. Palmer, A. Variable pricing through revenue management: a critical evaluation of affective outcomes / A. Palmer, U. McMahon-Beattie // *Management research news*. – 2008. – Vol. 31, № 3. – P. 189–199.
24. Besanko, D. Optimal price skimming by a monopolist facing rational consumers / D. Besanko, W.L. Winston // *Management science*. – 1990. – Vol. 36, № 5. – P. 555–567.
25. Anderson, C.K. Wait or Buy? The strategic consumer: pricing and profit implications / C.K. Anderson, J.G. Wilson // *The Journal of the Operational Research Society*. – 2003. – Vol. 54. – P. 299–306.
26. Interval scheduling: a survey / A.W.J. Kolen [et al.] // *Naval Research Logistics*. – 2007. – Vol. 54, № 5. – P. 530–543.
27. Kovalyov, M.Y. Fixed interval scheduling: models, applications, computational complexity and algorithms / M.Y. Kovalyov, C.T. Ng, T.C.E. Cheng // *European Journal of Operational Research*. – 2007. – Vol. 178. – P. 331–342.
28. A graph-theoretic approach to interval scheduling on dedicated unrelated parallel machines / C.T. Ng [et al.] // *Journal of the Operational Research Society* [Electronic resource]. – Mode of access : <http://goo.gl/kOW6fH>. – Date of access : 22.04.2014.
29. Pullman, M. Capacity management for hospitality and tourism : a review of current approaches / M. Pullman, S. Rogers // *International Journal of Hospitality Management*. – 2010. – Vol. 29, № 1. – P. 177–187.
30. Hadjinicola, G.C. The overbooking problem in hotels with multiple tour-operators / G.C. Hadjinicola, C. Panayi // *International Journal of Operations and Production Management*. – 1997. – Vol. 17, № 9. – P. 874–885.
31. Ivanov, S. Dynamic overbooking limits for guaranteed and nonguaranteed hotel reservations / S. Ivanov // *Tourism Today*. – 2007. – Vol. 7. – P. 100–108.
32. Ivanov, S. Management of overbookings in the hotel industry – basic concepts and practical challenges / S. Ivanov // *Tourism Today*. – 2006. – Vol. 6. – P. 19–32.
33. Koide, T. The hotel yield management with two types of room prices, overbooking and cancellations / T. Koide, H. Ishii // *International Journal of Production Economics*. – 2005. – Vol. 93, № 94. – P. 417–428.
34. Netessine, S. Introduction to the theory and practice of yield management / S. Netessine, R. Shumsky // *INFORMS Transactions on Education*. – 2002. – Vol. 3, № 1. – P. 34–44.
35. Kimes, S.E. The strategic levers of yield management / S.E. Kimes, R.B. Chase // *Journal of Service Research*. – 1998. – Vol. 1, № 2. – P. 156–166.
36. Vinod, B. Unlocking the value of revenue management in the hotel industry / B. Vinod // *Journal of Revenue and Pricing Management*. – 2004. – Vol. 3, № 2. – P. 178–190.
37. Hanks, R.D. Discounting in the hotel industry. A new approach / R.D. Hanks, R.G. Cross, R.P. Noland // *Cornell Hotel and Restaurant Administration Quarterly*. – 2002. – Vol. 43, № 4. – P. 94–103.
38. Ng, I.C.L. A demand-based model for the advance and spot pricing of services / I.C.L. Ng // *Journal of Product and Brand Management*. – 2009. – Vol. 18, № 7. – P. 517–528.
39. Zhang, M. Fencing in the context of revenue management / M. Zhang, P.C. Bell // *International Journal of Revenue Management*. – 2010. – Vol. 4, № 1. – P. 42–68.

40. Carvell, S.A. Exotic reservations – low price guarantee / S.A. Carvell, D.A. Quan // *International journal of hospitality management*. – 2008. – Vol. 27, № 2. – P. 162–169.
41. Room rate parity analysis across different hotel distribution channels in the U.S. / T. Demirciftci [et al.] // *Journal of Hospitality Marketing & Management*. – 2010. – Vol. 19, № 4. – P. 295–308.
42. Bitran, G. Periodic pricing of seasonal product in retailing / G. Bitran, S. Mondschein // *Management science*. – 1997. – Vol. 43. – P. 427–443.
43. Feng, Y. Optimal starting times for end-of-season sales and optimal stopping times for promotional fares / Y. Feng, G. Gallego // *Management science*. – 1995. – Vol. 41. – P. 1371–1391.
44. Gallego, G. Optimal dynamic pricing of inventories with stochastic demand over finite horizons / G. Gallego, G. van Ryzin // *Management science*. – 1994. – Vol. 40. – P. 999–1020.
45. Brynjolfsson, E. Frictionless Commerce? A Comparison of internet and conventional retailers / E. Brynjolfsson, M. Smith // *Management science*. – 1999. – Vol. 46, № 4. – P. 563–585.
46. Boyd, E.A. Revenue management and e-commerce / E.A. Boyd, I.C. Bilegan // *Management science*. – 2003. – Vol. 49, № 10. – P. 1363–1386.
47. van Ryzin, G. A multi-product dynamic pricing problem and its applications to network yield management / G. van Ryzin, G. Gallego // *Operations research*. – 1997. – Vol. 45. – P. 24–41.
48. Ladany, S. Optimal cruise-liner passenger cabin pricing policy / S. Ladany, A. Arbel // *European journal of operational research*. – 1991. – Vol. 55, № 2. – P. 136–147.
49. You, P.S. Dynamic pricing in airline seat management for flights with multiple flight legs / P.S. You // *Transportation science*. – 1999. – Vol. 33, № 2. – P. 192–206.
50. Gaimon, C. Simultaneous and dynamic price, production, inventory and capacity decisions / C. Gaimon // *European journal of operational research*. – 1988. – Vol. 35. – P. 426–441.
51. Lau, A.H.L. The newsboy problem with price dependent demand distribution / A.H.L. Lau, H.S. Lau // *IIE transactions*. – 1988. – Vol. 20, № 2. – P. 168–175.
52. Weatherford, L.R. Optimization of perishable-asset revenue management problems that allow prices as decision variables / L.R. Weatherford // *International journal of services technology and management*. – 2001. – Vol. 2, № 1/2. – P. 71–101.
53. Chatwin, R.E. Optimal dynamic pricing of perishable products with stochastic demand and a finite set of prices / R.E. Chatwin // *European journal of operational research*. – 2000. – Vol. 125. – P. 149–174.
54. Feng, Y. Perishable asset revenue management with Markovian time dependent demand intensities / Y. Feng, G. Gallego // *Management science*. – 2000. – Vol. 46. – P. 941–956.
55. Feng, Y. A continuous-time yield management model with multiple prices and reversible price changes / Y. Feng, B. Xiao // *Management science*. – 2000. – Vol. 48. – P. 644–657.
56. Feng, Y. Optimal policies of yield management with multiple predetermined prices / Y. Feng, B. Xiao // *Management science*. – 2000. – Vol. 48. – P. 332–343.
57. Tranter, K.A. Introduction to revenue management for the hospitality industry / K.A. Tranter, T. Stuart-Hill, J. Parker. – Harlow : Pearson Prentice hall, 2008. – 352 p.
58. Emeksiz, M. A yield management model for five-star hotels: computerized and non-computerized implementation / M. Emeksiz, D. Gursoy, O. Icoz // *International journal of hospitality management*. – 2006. – Vol. 25, № 4. – P. 536–551.
59. Cross, R.G. Revenue management / R.G. Cross. – N.Y. : Broadway books, 1997. – 288 p.
60. Pölt, S. Forecasting is difficult – especially if it refers to the future / S. Pölt // *Reservations and yield management study group annual meeting proceedings*. – Melbourne, 1998. – 32 p.
61. Weatherford, L.R. Forecasting for hotel revenue management: testing aggregation against disaggregation / L.R. Weatherford, S.E. Kimes, D.A. Scott // *Cornell hotel and restaurant administration quarterly*. – 2001. – Vol. 42. – P. 53–64.
62. Lee, A.O. Airline reservations forecasting: probabilistic and statistical models of the booking process: PhD thesis / O.A. Lee. – Cambridge, 1990. – 266 p.
63. The accuracy of extrapolation (time series) methods: results of a forecasting competition / S. Makridakis [et al.] // *Journal of forecasting*. – 1982. – Vol. 1. – P. 111–153.
64. Weatherford, L.R. A comparison of forecasting methods for hotel revenue management / L.R. Weatherford, S.E. Kimes // *International journal of forecasting*. – 2003. – Vol. 19. – P. 401–415.

65. Fildes, R. Forecasting competitions – their role in improving forecasting practice and research / R. Fildes, K. Ord // *A companion to economic forecasting* / M.P. Clements, D.F. Hendry. – Oxford, 2002. – Ch. 15. – P. 322–353.
66. Ben-Akiva, M. Improving airline passenger forecasts using reservation data / M. Ben-Akiva // *Fall ORSA/TIMS conference*. – St. Louis, 1987.
67. Forecasting uncertain hotel room demand / M. Rajopadhye [et al.] // *Information sciences*. – 2001. – Vol. 132, № 1–4. – P. 1–11.
68. Yüksel, S. An integrated forecasting approach to hotel demand / S. Yüksel // *Mathematical and computer modelling*. – 2007. – Vol. 46, № 7, 8. – P. 1063–1070.
69. Lim, C. An econometric analysis of hotel-motel room nights in New Zealand with stochastic seasonality / C. Lim, F. Chan // *International journal of revenue management*. – 2011. – Vol. 5, № 1. – P. 63–83.
70. Lim, C. Forecasting hotel guest nights in New Zealand / C. Lim, C. Chang, M. McAleer // *International journal of hospitality management*. – 2009. – Vol. 28, № 2. – P. 228–235.
71. Armstrong, J.S. Error measures for generalizing about forecasting methods: empirical comparisons / J.S. Armstrong, F. Collopy // *International journal of forecasting*. – 1992. – Vol. 8. – P. 69–80.
72. Zakhary, A. A comparative study of the pickup method and its variations using a simulated hotel reservation data / A. Zakhary, N. El. Gayar, A.F. Atiya // *ICGST international journal on artificial intelligence and machine learning*. – 2008. – Vol. 8. – P. 15–21.
73. Schnaars, S.P. Situational factors affecting forecast accuracy / S.P. Schnaars // *Journal of marketing research*. – 1984. – Vol. 21. – P. 290–297.
74. Weatherford, L.R. A tutorial on optimization in the context of perishable-asset revenue management problems for the airline industry / L.R. Weatherford // *Operations research in the airline industry* / ed. G. Yu. – Boston, 1998. – P. 68–100.
75. Pak, K. Airline revenue management: an overview of OR techniques 1982–2001 / K. Pak, N. Piersma // *Erasmus University Rotterdam [Electronic resource]*. – 2002. – Mode of access : <http://repub.eur.nl/pub/584>. – Date of access : 22.04.2014.
76. Littlewood, K. Forecasting and control of passenger bookings / K. Littlewood // *AGIFORS symposium proc. 12, Nathanya, 1972; Alliance group of the international federation of operational research scientists*. – Nathanya, 1972.
77. Belobaba, P.P. Air travel demand and airline seat inventory management: PhD thesis / P.P. Belobaba. – Cambridge, 1987. – 236 p.
78. Curry, R.E. Optimal airline seat allocation with fare classes nested by origin and destinations / R.E. Curry // *Transportation science*. – 1990. – Vol. 24. – P. 193–204.
79. Wollmer, R.D. An airline seat management model for a single leg route when lower fare classes book first / R.D. Wollmer // *Operations research*. – 1992. – Vol. 40. – P. 26–37.
80. Brumelle, S.L. Airline seat allocation with multiple nested fare classes / S.L. Brumelle, J.I. McGill // *Operations research*. – 1993. – Vol. 41. – P. 127–137.
81. Lee, T.C. A model for dynamic airline seat inventory control with multiple seat bookings / T.C. Lee, M. Hersh // *Transportation science*. – 1993. – Vol. 27. – P. 252–265.
82. Kleywegt, A.J. The dynamic and stochastic knapsack problem / A.J. Kleywegt, J.D. Papastavrou // *Operations research*. – 1998. – Vol. 46. – P. 17–35.
83. Subramanian, J. Airline yield management with overbooking, cancellations, and no-shows / J. Subramanian, Jr. S. Stidham, C.J. Lautenbacher // *Transportation science*. – 1999. – Vol. 33. – P. 147–167.
84. Williams, H.P. Model building in mathematical programming / H.P. Williams. – Chichester, N.Y. : Wiley, 1999. – 353 p.
85. Писарук, Н.Н. Модели и методы смешанно-целочисленного программирования / Н.Н. Писарук. – Минск : БГУ, 2010. – С. 205–208.
86. The passenger mix problem in the scheduled airlines / F. Glover [et al.] // *Interfaces*. – 1982. – Vol. 12. – P. 73–79.
87. Talluri, K.T. A randomized linear programming method for computing network bid prices / K.T. Talluri, G.J. van Ryzin // *Transportation science*. – 1999. – Vol. 33. – P. 207–216.

88. Williamson, E.L. Airline network seat inventory control: methodologies and revenue impacts: PhD thesis / E.L. Williamson. – Cambridge, 1992. – 270 p.
89. Wei, Y.J. Airline O-D control using network displacement concepts: MS thesis / Y.J. Wei. – Cambridge, 1997. – 139 p.
90. Talluri, K.T. An analysis of bid-price controls for network revenue management / K.T. Talluri, G.J. van Ryzin // *Management science*. – 1998. – Vol. 44, № 11. – P. 1577–1593.
91. Chen, V.C.P. A Markov decision problem based approach to the airline YM problem / V.C.P. Chen, D. Gunther, E.L. Johnson // Georgia institute of technology, The Logistics Institute [Electronic resource]. – Mode of access : <http://goo.gl/H77hjL>. – Date of access : 22.04.2014.
92. van Ryzin, G. Revenue management / G. van Ryzin, K.T. Talluri // *Handbook of transportation science* / ed. R.W. Hall. – Boston, 2003. – Ch. 16. – P. 599–659.
93. Cooper, W.L. A class of hybrid methods for revenue management / W.L. Cooper, T. Homem-de-Mello // Northwestern university, Department of industrial engineering and management sciences [Electronic resource]. – Mode of access : <http://goo.gl/pJzeF9>. – Date of access : 22.04.2014.
94. Lai, K.-K. A stochastic approach to hotel revenue optimization / K.-K. Lai, W.-L. Ng // *computers and operations research*. – 2005. – Vol. 32, № 5. – P. 1059–1072.
95. A stochastic approach to hotel revenue management considering multiple-day stays / S. Liu [et al.] // *International journal of information technology and decision making*. – 2006. – Vol. 5, № 3. – P. 545–556.
96. Liu, S. Booking models for hotel revenue management considering multiple-day stays / S. Liu, K.K. Lai, S.-Y. Wang // *International journal of revenue management*. – 2008. – Vol. 2, № 1. – P. 78–91.
97. Baker, T.K. The benefits of optimizing prices to manage demand in hotel revenue management systems / T.K. Baker, D.A. Collier // *Production and operations management*. – 2003. – Vol. 12. – P. 502–518.
98. Forecasting hotel arrivals and occupancy using Monte Carlo simulation / A. Zakhary [et al.] // *Journal of revenue and pricing management*. – 2011. – Vol. 10, № 4. – P. 344–366.
99. Bertsimas, D. Restaurant revenue management / D. Bertsimas, R. Shioda // *Operations research*. – 2003. – Vol. 51. – P. 472–486.
100. Badinelli, R.D. An optimal, dynamic policy for hotel yield management / R.D. Badinelli // *European journal of operations research*. – 2000. – Vol. 121. – P. 476–503.
101. Padhi, S.S. Competitive revenue management for fixing quota and price of hotel commodities under uncertainty / S.S. Padhi, V. Aggarwal // *International journal of hospitality management*. – 2011. – Vol. 30, № 3. – P. 725–734.
102. Dynamic pricing and the direct-to-customer model in the automotive industry / S. Biller [et al.] // *Electronic commerce research*. – 2005. – Vol. 5, № 2. – P. 309–334.

Поступила 27.11.2013

*Объединенный институт проблем
информатики НАН Беларуси,
Минск, Сурганова, 6
e-mail: andrei.bandalouski@gmail.com*

A.M. Bandalouski

AN OVERVIEW OF THE HOTEL REVENUE MANAGEMENT MODELS

The paper gives a brief scope of Hotel Revenue Management theory and discusses its basic definitions. It also suggests a new classification of Revenue Management processes and surveys dynamic pricing, forecasting and optimization models employed in Hotel Revenue Management. Last section proposes promising directions of future research.

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.67, 575.112:004

А.В. Саечников¹, Н.Н. Яцков¹, П.В. Назаров², Л. Валлар², В.В. Апанасович¹**АНАЛИЗ ЭКСПРЕССИИ ГЕНОВ В РЕЗУЛЬТАТЕ ВОЗДЕЙСТВИЯ
ИНТЕРФЕРОНА IFN- γ НА КЛЕТКУ С ИСПОЛЬЗОВАНИЕМ
ПРОГРАММНОГО ПАКЕТА GeneExpressionAnalyser**

Предлагается программный пакет GeneExpressionAnalyser для анализа данных, полученных в ходе проведения экспериментов с использованием биочипов ДНК. Детально рассмотрены алгоритмы предварительного анализа данных, выделения дифференциально-выраженных генов и анализа биологических функций клетки. Работоспособность пакета исследуется на примере опубликованных экспериментальных данных, представляющих результаты эксперимента по исследованию изменений экспрессии генов в клетке меланомы под воздействием интерферона IFN- γ с течением времени.

Введение

Развитие биотехнологий напрямую связано с разработкой эффективных методов и алгоритмов обработки большого объема информации, получаемой от биологических микрочипов различного назначения. Олигонуклеотидные микрочипы ДНК позволяют изучать экспрессию генов [1]. Анализ данных, полученных в ходе проведения экспериментов с использованием микрочипов ДНК, дает возможность выявить функционально связанные и несвязанные гены, определить доминирующие биофункции клетки. Реализация данной задачи требует значительной рутинной обработки и количественной интерпретации экспериментальных данных, однако алгоритмические возможности для их реализации лишь ограниченно предложены в стандартных открытых статистических пакетах обработки данных [2, 3]. Практически каждый из пакетов имеет определенные недостатки, например не позволяет исследовать данные с наличием пропущенных значений, дубликатов, выбросов; содержит ограниченный круг алгоритмов кластерного анализа экспрессии генов и их аннотаций [4, 5]. В экспериментах с микрочипами ДНК даже результаты предварительной обработки изображений зависят от типа микрочипа. Это условие накладывает ограничения на типы принимаемых программным обеспечением данных и дальнейшую обработку результатов. Существует множество разработанных программных пакетов, как лицензионных, так и находящихся в свободном доступе. Во многих пакетах, таких как BETR, CAGED, GeneNetWeaver, SNOMAD, GeneMAPP, реализован только определенный алгоритм анализа данных [6–8]. Достойной альтернативной платформой является среда статистического программирования R [9]. Среда R имеет ряд преимуществ: используется в открытом и бесплатном проекте Bioconductor для программных разработок, включает широкий набор статистических функций, в том числе функций для анализа биочипов ДНК [10–12]. Однако проект предоставляет лишь ограниченный набор программных средств для создания графических интерфейсов (например, tcltk, RGtk2, gpanel, gWidgets, Shiny), облегчающих работу пользователя, что в конечном счете автоматически требует от пользователя знания языка программирования R или хотя бы частичного понимания программного кода.

Существуют и другие пакеты программного обеспечения для комплексного анализа данных, получаемых с помощью биочипов ДНК, например Coral [13], DMET-Analyser [14], eXframe [15]. Некоторые из пакетов являются бесплатными при условии дополнительного использования сторонних коммерческих программ, другие позволяют работать только со специализированными типами микрочипов. Программные пакеты Partek [16] и IPA [17] выполняют определенные этапы полного анализа данных. Поэтому их приходится использовать совместно

с другими программными продуктами, что требует дополнительных затрат на проведение работ по совместимости результатов входных и выходных данных. Пакет GoMiner [18] – мощный инструмент для выделения активных функций клетки, однако он имеет ряд недостатков. Во-первых, для получения списка значимых функций требуется использование дополнительных программных пакетов обработки данных и предоставление списка значимых генов. Во-вторых, архитектура пакета GoMiner основана на использовании интернет-подключения к удаленным базам данных, что приводит к значительному замедлению вычислений.

Целью данной работы является разработка программного комплекса для анализа широкого набора биочипов ДНК, интегрирующего основные этапы анализа данных: загрузку данных, предварительную обработку, выделение значимых генов, определение доминирующих функций клетки. Для всестороннего исследования разработанного программного пакета используются опубликованные наборы данных, полученные в ходе проведения эксперимента по исследованию воздействия интерферона IFN- γ на живую клетку человека [19].

1. Экспериментальные данные

Для демонстрации работоспособности пакета *GeneExpressionAnalyser* используются опубликованные экспериментальные данные [19], размещенные в хранилище ArrayExpress (индекс E-MEXP-3720). В работе изучается влияние интерферона IFN- γ на клеточную линию A375 меланомы. Клетки помещались в пробирки с питательной средой и культивировались в течение 96 ч. В определенные моменты времени клетки однократно подвергались воздействию IFN- γ (в концентрации 50 нг/мл). В качестве контрольных использовались клетки, не подвергавшиеся воздействию интерферона, а также клетки, в которых сигнальная система, состоящая из янус-киназы сигнального белка-трансдуктора и активатора транскрипции (JAK – STAT), заблокирована с помощью ингибитора янус-киназы I (JAK inhibitor I). Таким образом, были задействованы три биологических репликанта: клетки в отсутствие воздействия IFN- γ (ctrl), клетки с заблокированным сигнальным путем JAK – STAT (JI ctrl), а также культивированные клетки – в течение 3, 12, 24, 48 и 72 ч (далее моменты времени обозначены как 03Н, 12Н, 24Н, 48Н, 72Н) после добавления IFN- γ . В точке нулевого момента времени рассматривается контрольный образец ctrl, а в точке 72 ч – контрольный образец JI ctrl. Для временных точек 03Н, 12Н проведены по два эксперимента с биочипами, для временных точек 24Н, 48Н и 72Н – по три эксперимента для каждого образца. Из клеточного материала выделена РНК, которая после соответствующей обработки была помещена на микрочипы Affymetrix Gene Chip Human Gene 1.0 ST Array. Детальное описание эксперимента и его результатов приведено в [19]. Выполнен анализ 17 биочипов, представляющих временные точки и различные технические репликаны. На каждом биочипе размещено по 33 252 гена.

2. Программный пакет *GeneExpressionAnalyser*

В качестве основы среды разработки и реализации пакета выбрана программная среда Matlab с библиотекой Bioinformatics. Для построения графического интерфейса использовалась система GUIDE пакета MATLAB 7.11.0 (R2010b) для ОС Windows®. Среда Matlab выбрана по следующим причинам: 1) из-за возможности разработки хорошо зарекомендованных standalone-приложений, дружественного графического интерфейса и работы в режиме офлайн (достаточно нескольких подключений для обновления базы данных биофункций), что, несомненно, приставляет интерес для работы пользователей из стран СНГ; 2) надежности и оптимизации ядра среды, реализованного на языке программирования C++, для реализаций операций математических вычислений с матрицами, что значительно увеличивает скорость анализа и моделирования больших объемов многомерных данных; 3) наличия широкого набора отлаженных функций и программных инструментов для анализа биочипов ДНК, часть из которых существенно программно оптимизирована (что увеличивает эффективность и точность анализа). Разработанный пакет является бесплатным при условии распространения скомпилированного исполнительного файла с набором необходимых для работы библиотек.

Также бесплатно распространяется файл среды Matlab при условии наличия предварительно установленного пакета Matlab на компьютере-клиенте.

GeneExpressionAnalyser осуществляет следующие функции (рис. 1):

- загрузку и предварительную фильтрацию некачественных данных;
- предварительную обработку данных (опционально);
- нормировку данных [20–23];
- восстановление пропущенных значений [24];
- поиск дифференциально-выраженных генов (с использованием метода анализа значимости биочипов, далее используется англоязычная аббревиатура SAM (от Significance Analysis of Microarrays) [25]);
- иерархическую и неиерархическую кластеризацию [26];
- визуализацию данных с использованием метода главных компонент [27];
- выделение статистически значимых функций в ходе анализа аннотаций генов (англ. GeneOntology-анализ, GO [28]) с применением методов Фишера [18] и случайных перестановок [29].

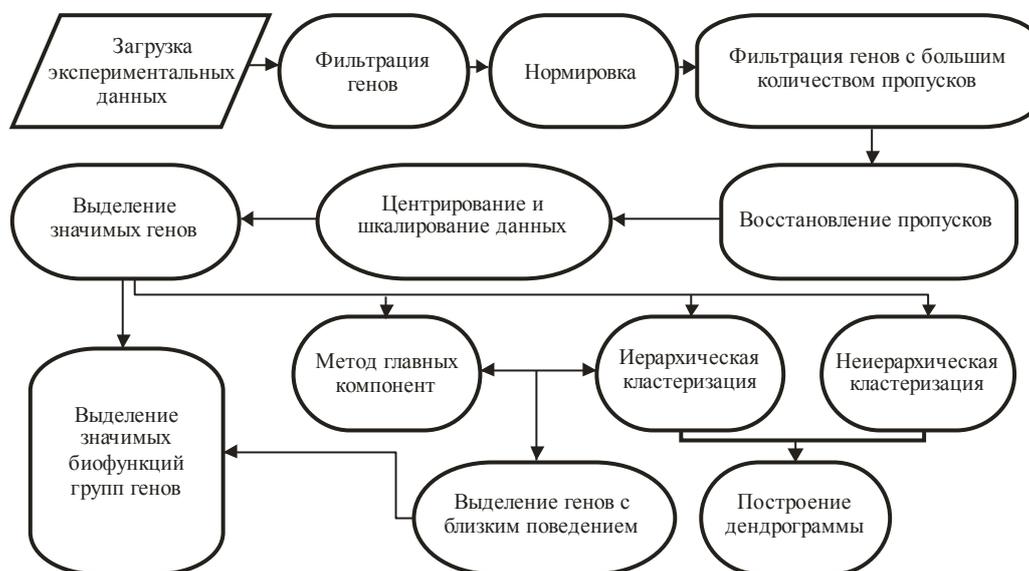


Рис. 1. Функциональная схема обработки данных с использованием GeneExpressionAnalyser

Рассмотрим подробно этапы анализа данных.

Загрузка данных. В программе предлагается выбрать один из трех вариантов формата загрузки данных:

1. Формат данных двухканальных микрочипов (*.grg-файлы) [30]. После загрузки производится удаление из рассмотрения данных, для которых значение параметра $Flag = 0$ [30].

2. Формат данных чипов Affymetrix [31]. В этом случае требуется загрузка *.cel-файлов, которые содержат информацию об интенсивности люминесценции ячеек микрочипа, а также *.cdf-библиотеку, в которой интегрирована информация о ДНК-мишенях микрочипа. После загрузки предлагается выполнить фильтрацию генов с низким значением интенсивности экспрессии.

3. Загрузка данных из таблиц Excel, содержащих MA-значения [32] или непосредственно значения экспрессии генов [33].

Предварительная обработка данных (на примере двухканальных микрочипов). Процедура включает вычитание шума в каналах, определение уровня экспрессии и средней интенсивности в каналах:

$$M = \log_2 \left(\frac{R-Rb}{G-Gb} \right), \quad A = \frac{1}{2} \log_2 ((R-Rb)(G-Gb)),$$

где M – оценка уровня экспрессии гена; A – значение средней интенсивности флуоресценции сигнала гена в красном и зеленом каналах для некоторой ячейки биочипа; R, G – значения интенсивности ячейки биочипа в красном и зеленом каналах соответственно; Rb, Gb – фоновые значения интенсивности ячейки в красном и зеленом каналах соответственно.

Для чипов Affymetrix данная процедура не проводится, так как данные представляют собой логарифм интенсивности, пропорциональный логарифму концентрации РНК.

Нормировка. Для данных, снятых с помощью двухцветных матриц, применяется нормировка методом локального сглаживания графиков разброса (locally weighted scatter plot smoothing [20]). Этот шаг необходим для снижения влияния факторов, непосредственно связанных с конкретным биочипом. Для Affymetrix-данных производится RMA(Robust Multichip Average)-нормировка [21, 22] по каждому из чипов.

RMA-нормировка состоит из трех этапов. На первом этапе производится расчет фоновой компоненты для каждого биочипа в отдельности. (Предполагается, что фоновая компонента шума подчиняется нормальному распределению, а сигнальная – экспоненциальному распределению.) Фоновая компонента вычитается из набора данных.

На втором этапе производится квантильная нормировка по всему набору значений, в результате которой выравниваются эмпирические распределения интенсивностей проб для каждого биочипа. На последнем этапе производится суммирование интенсивностей проб в проб-сету с использованием метода медианной очистки (Median Polish [21, 22]).

Корректировка матрицы уровней экспрессии. Данная процедура связана с устранением пропущенных значений в таблице исходных данных. Пропущенные значения являются результатом повреждения биочипа или неполной гибридизации биологического материала. Если у гена пропущено больше некоторого определенного процента значений экспрессии (как правило, 33 % значений) в сводной таблице экспериментальных данных, то в дальнейшем данный ген исключается из последующего анализа. В разработанном пакете реализована процедура восстановления пропущенных значений. Для восстановления пропущенных значений используются метод k -ближайших соседей и его различные модификации [24]. Эффективность и качество методов восстановления пропущенных значений падают в случае анализа времязависимых данных [24]. Как правило, для чипов Affymetrix данная процедура не используется в силу высокого качества изготовления чипов, высокой плотности записи и точности.

В ходе выполнения данного этапа производится центрирование и шкалирование данных с целью устранения неоднородности в данных:

$$M_{ij} = \frac{x_{ij} - \bar{M}_j}{\delta(\bar{M}_j)},$$

где \bar{M}_j и $\delta(\bar{M}_j)$ – математическое ожидание и среднеквадратическое отклонение по j -й колонке (биочип); M_{ij} – значение экспрессии i -го гена на j -м биочипе. Данная нормировка проводится по среднему всех значений экспрессии генов по каждому из биочипов в отдельности.

Для дальнейшего анализа производится отбор наиболее информативных данных. Простейший способ состоит в исключении генов с низкими уровнями экспрессии. Пороговое значение выбирается произвольно или экспертным решением. Производится группировка генов (усреднение значений экспрессии для экземпляров одного гена).

Выделение значимых генов методом SAM [25]. Метод SAM основан на анализе флуктуаций в экспериментальных данных. Для определения неслучайных флуктуаций вводится статистика

$$d_i = \frac{r_i}{s_i + s_0},$$

где r_i – нормированное значение экспрессии i -го гена; s_i – стандартное отклонение значения экспрессии гена; s_0 – малая постоянная величина, введенная для уменьшения зависимости d_i от уровня экспрессии гена.

Величина s_0 рассчитывается итерационно на основе значений r_i и s_i . Производится упорядочивание значений $d(i)$ в порядке убывания. Для каждого гена из набора случайно сгенерированных данных рассчитывается d_{pi} , затем гены выстраиваются в порядке убывания значений d_{pi} . Ожидаемое относительное различие в уровнях экспрессии d_e определяется как среднее по всем случайно сгенерированным данным. Следующий этап – построение диаграммы рассеяния наблюдаемого относительного различия от ожидаемого относительного различия в уровнях экспрессии. Гены, которые расположены на диаграмме рассеяния на расстоянии, большем чем Δ (данный параметр подбирается) от диагональной прямой, считаются значимыми. Оценочное значение ошибочно найденных значимых генов (false discovery rate – FDR) рассчитывается как среднее значение числа генов, названных значимыми для всех перестановок. Варианты метода SAM различаются в методике расчета величин r_i и s_i . Например, для однофакторного анализа значимости (SAM one class)

$$r_i = x_i = \sum_j \frac{x_{ij}}{n}, \quad s_i = \sqrt{\sum_j \frac{(x_{ij} - x_i)^2}{n(n-1)}},$$

где j – номер экспериментального значения для одного гена; x_i – среднее значение экспрессии гена; x_{ij} – значение экспрессии i -го гена в j -м биочипе; n – количество биочипов. Для многофакторного анализа значимости (SAM-multiclass)

$$x_{im} = \sum_j \frac{x_{ij}^m}{n};$$

$$r_i = \sqrt{\left(\frac{\sum n_m}{\prod n_m}\right) \sum_{m=1}^C n_m (x_{im} - x_i)^2};$$

$$s_i = \sqrt{\left(\frac{1}{\sum (n_m - 1)}\right) \left(\sum \frac{1}{n_m}\right) \sum_{m=1}^C \sum_{j \in C_m} (x_{ij} - x_{im})^2},$$

где C – количество классов данных; C_m – индекс наблюдений в классе данных m ; n_m – количество наблюдений в классе m . Подробная информация по методике расчета SAM представлена в работе [29].

В программном пакете реализованы различные варианты анализа методом SAM (однофакторный, двухфакторный парный (непарный) и временной анализы). Метод SAM реализован на базе опубликованной R -функции [34]. В результате анализа данных с использованием метода SAM формируются списки дифференциально выраженных генов.

Кластеризация. Профили экспрессии отражают изменение транскрипции в зависимости от внешних условий или концентрации веществ, поэтому в пакете GeneExpressionAnalyser реализована возможность классификации генов и биочипов. Реализованы методы неиерархического (метод k -средних [26]) и иерархического кластерного анализа [27, 35]. В результате кластерного анализа формируются основные профили генной экспрессии.

Анализ данных с использованием метода главных компонент. Метод главных компонент [36] реализован для оценки качества кластеризации данных, визуализации данных в пространстве низкой размерности, определения выбросов в пакете. Характерной особенностью метода являются выбор и оценка значимости не отдельных переменных, а информативных по совокупности групп переменных. Пользователю предоставляется возможность построения графиков разброса в пространстве различных главных компонент (рис. 2).

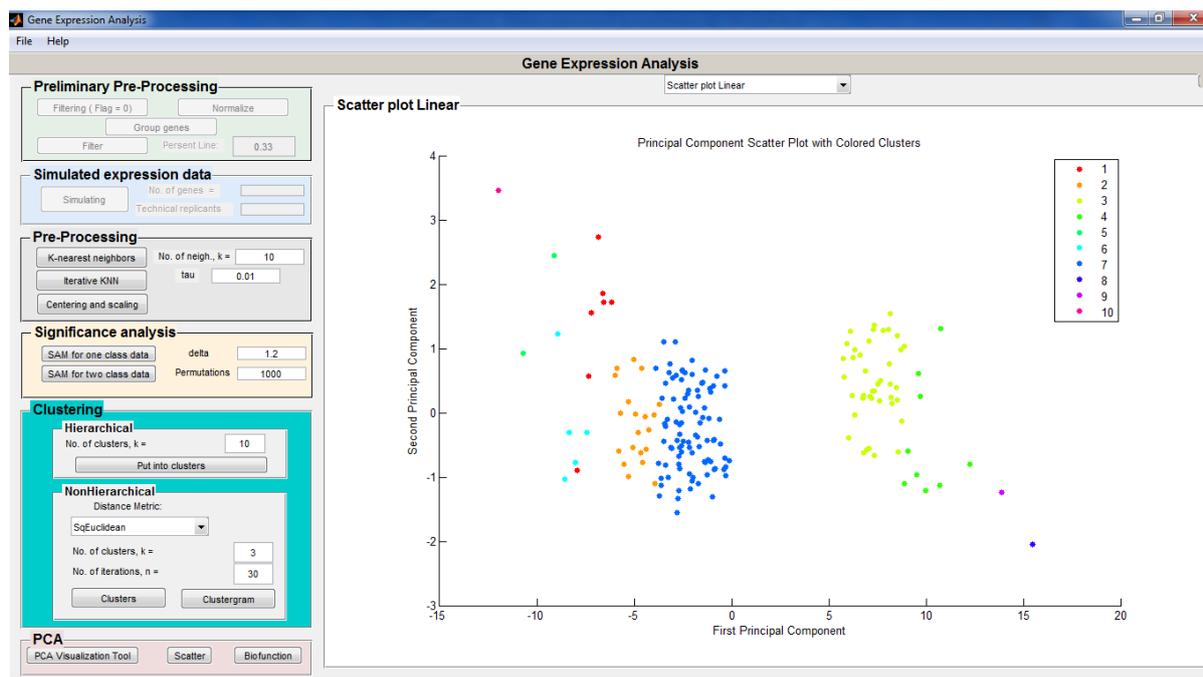


Рис. 2. Рабочее окно GeneExpressionAnalyser

Выделение значимых функций генов в ходе анализа генных аннотаций. Данный этап достаточно полно реализован в программном пакете GoMiner [18, 29], а также в программном пакете IPA® (Ingenuity Pathway Analysis). За алгоритмическую основу в разработанном пакете взяты (и усовершенствованы) подходы, используемые в названных выше пакетах. При этом оставлена возможность подключения к удаленным GO-базам данных. Достоинством GeneExpressionAnalyser является возможность определения значимых аннотаций для подгрупп дифференциально выраженных генов. В программе реализован метод одностороннего точного критерия Фишера [27]. Метод используется для определения статистической значимости категорий аннотаций онтологии генов, характеризуемой p -величиной [27]. Пользователь выбирает требуемый уровень p -значения (например, $p = 0,05$), после чего формируется список GO-аннотаций.

Большинство ключевых параметров анализа вводятся пользователем с помощью стандартного оконного интерфейса ОС Windows® (см. рис. 2). Вывод промежуточных и итоговых результатов осуществляется в виде графиков и внешних объектов (баз данных для возможности анализа результатов, итоговых таблиц, дендрограмм). Предусмотрена возможность сохранения результатов анализа данных в графические файлы. Промежуточные и конечные результаты анализа можно сохранить в специальном формате с возможностью последующего открытия и исследования. Работоспособность отдельных составляющих программного пакета подтверждена и исследована на примерах смоделированных и опубликованных экспериментальных данных [37, 38].

3. Имитационное моделирование

В GeneExpressionAnalyser интегрирована имитационная модель, генерирующая выражения экспрессии генов в эксперименте с биочипами ДНК. Имитационная модель предназначена с целью оценки устойчивости и проверки работоспособности разрабатываемых алгоритмов пакета, а также моделирования пользователем различных условий и возможностей проведения экспериментов с биочипами ДНК. Модель воспроизводит матрицы уровней экспрессии генов. Этапы обработки данных до этапа центрирования и шкалирования не учитываются. В модели задается вид профиля экспрессии группы генов, затем накладывается шум [39]. Добавление гауссова шума к профилю экспрессии производится следующим образом:

$$M_{sim} = M_{def} + \sigma * z,$$

где M_{def} – априорный профиль экспрессии в логарифмической шкале; σ – среднеквадратическое отклонение шума (выбирается исходя из оценки параметров распределения шума, наблюдаемого в конкретном эксперименте); z – реализация случайной величины со стандартным нормальным распределением $N(0,1)$.

Модель способна генерировать данные времязависимых экспериментов, для чего задается изменяющийся во времени априорный профиль экспрессии, а также биочипы размером до 50 000 ячеек. Количество дифференциально выраженных генов не должно превышать 20 % от общего числа генов. Ограничение обусловлено трудоемкостью вычислений при проведении многофакторного и однофакторного временного SAM-анализов, так как для работы данного метода необходимо одновременно работать с данными большого количества биочипов. В работе производилось моделирование данных для проведения проверки работоспособности программного обеспечения в целом и отдельных методов, входящих в состав пакета. Модель можно применять для моделирования значений экспрессии РМ(Perfect Match)-проб генов (25-мерных последовательностей, комплементарных эталонной на биочипе). В данном варианте моделирования значения экспрессии характеризуются экспоненциальной плотностью распределения.

4. Результаты исследования

В ходе проверки работоспособности пакета GeneExpressionAnalyser выполнен анализ 17 наборов экспериментальных данных. Рассмотрим основные результаты, полученные на каждом из этапов анализа. Ввиду высокого качества данных фильтрация с большим количеством пропусков не производилась. Выполнена RMA-нормализация набора данных. На первом этапе нормировки определены плотности интенсивности фоновой компоненты для каждого биочипа, относительно которых в дальнейшем были отнормированы значения экспрессии генов.

На рис. 3. показано распределение значений интенсивности $\log_2(\text{PM})$ [33], полученной после предварительной обработки изображений микрочипов. В результате нормировки преобразованы матрицы уровней экспрессии для каждого технического репликанта.

На этапе выделения дифференциально значимых генов выполнены следующие варианты метода SAM ($\text{FDR} < 0,05$): однофакторный, двухфакторный парный, двухфакторный непарный, однофакторный временной, многофакторный.

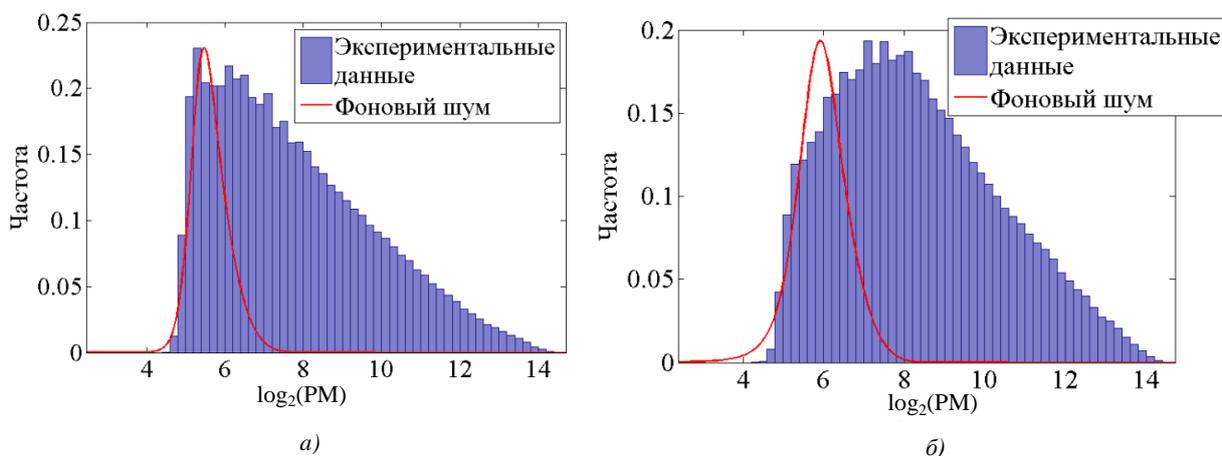


Рис. 3. Гистограммы распределения интенсивности $\log_2(\text{PM})$ и рассчитанной плотности фоновой компоненты:
 а) первый технический репликант в момент времени 12 ч после начала воздействия;
 б) второй технический репликант в момент времени 48 ч после начала воздействия

Однофакторный, двухфакторный парный, двухфакторный непарный SAM-методы использовались для выделения значимых генов в рамках каждой временной точки эксперимента, однофакторный временной и многофакторный SAM-методы – для получения списка значимых генов по всему эксперименту. Для метода SAM one class значения экспрессии нормировались относительно среднего значения контроля, что давало возможность воспринимать данные в рамках одной временной точки как данные одного класса.

Для двухфакторного парного и двухфакторного непарного SAM-методов для каждой временной точки строились два класса данных: контрольные и подверженные влиянию IFN-γ. В случае парного метода были синхронизированы значения для различных микрочипов. Для микрочипа, представляющего третий контрольный образец (культивированные клетки после добавления IFN-γ), было сформировано контрольное значение как среднее между контрольными значениями первого и второго микрочипов). Конечной целью SAM-метода является выделение как можно большего числа значимых генов с сохранением пропорции ошибочных генов ниже заданного порога, поэтому исходя из полученных данных (рис. 4, а) было решено использовать результаты, полученные вариантами двухфакторного парного и однофакторного методов.

Результаты обработки методом SAM

Тип метода SAM	03H		12H		24H		48H		72H		JII ctrl	
	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg	Pos	Neg
Однофакторный	59	0	158	0	2634	5791	2103	6923	4740	8580	3	0
Двухфакторный парный	71	10	337	40	1549	3887	2510	6281	3361	5954	8	1
Двухфакторный непарный	60	0	427	0	772	1279	1030	3607	2306	4942	131	85
Однофакторный временной	Pos		Neg		FDR		Многофакторный		Significant		FDR	
	2673		481		0,038				1242		0,04	

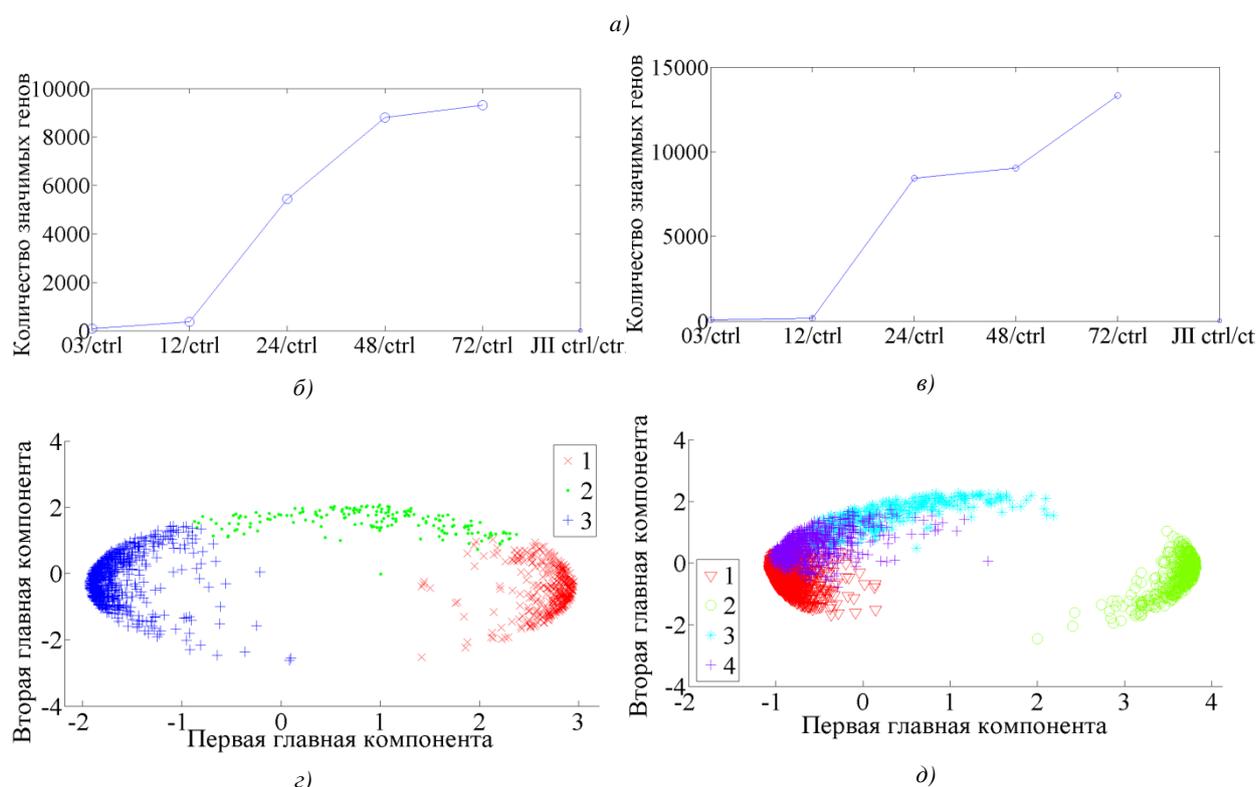


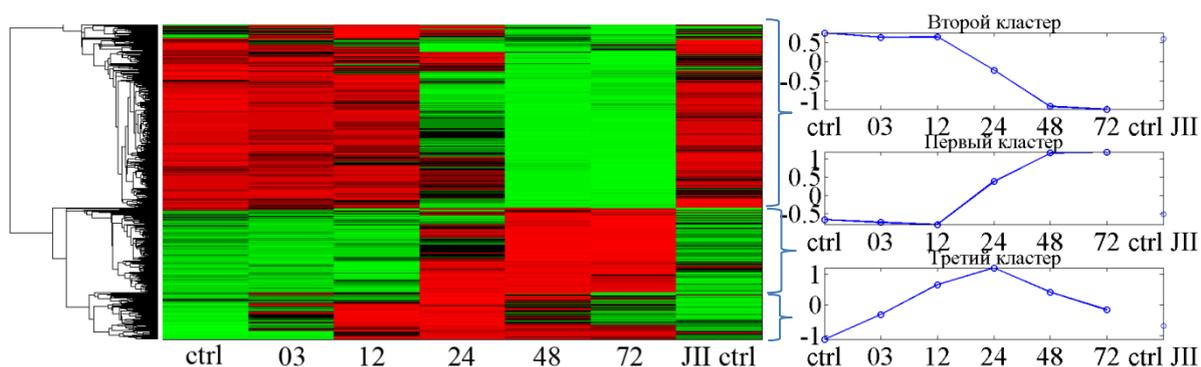
Рис. 4. Результаты дифференциального анализа с использованием SAM-метода: а) результат обработки SAM-методом (Pos – выразенные гены, Neg – подавленные гены); зависимости количества значимых генов от времени: б) SAM two class paired; в) SAM one class; графики разброса экспрессий генов в первых двух главных компонентах (маркерами обозначены различные кластеры генов): г) результаты для многофакторного SAM: первая компонента – 70,1 % дисперсии всех признаков; вторая – 15,4 % дисперсии всех признаков; д) результаты для однофакторного временного SAM: первая компонента – 62,7 % дисперсии всех признаков; вторая – 15,6 % дисперсии всех признаков

Необходимо отметить, что максимальный эффект воздействия на клетку IFN- γ в обоих случаях наблюдался в промежуток времени с 12 до 24 ч после начала воздействия. При этом в 24 ч около 70 % всех значимых генов являлись подавленными генами. Первая реакция клетки на IFN- γ наблюдается по прошествии определенного промежутка времени. Максимальный эффект воздействия интерферона падает после 48 ч (изменение количества значимых генов с 48 до 72 ч малó). Результаты однофакторного SAM-метода демонстрируют ослабление воздействия интерферона в промежуток времени с 24 до 48 ч и последующее его усиление (рис. 4, б, в). Несмотря на то что количество значимых генов в случае однофакторного анализа больше, для дальнейшего анализа оставлены результаты двухфакторного парного SAM-метода, так как сами входные данные больше подходят под определение двухфакторных данных: класс данных до воздействия и класс данных после воздействия. Следует отметить, что в случае двухфакторных данных производились наименьшие изменения матриц уровней экспрессии, что является дополнительным фактором в пользу выбора данного результата. В случае двухфакторного парного SAM-метода характер изменения количества значимых генов аналогичен характеру изменений, полученному с использованием пакета *limma* [19].

Для последующей кластеризации используются дифференциально выраженные гены, полученные многофакторным и однофакторным временным SAM-методами. В ходе анализа результатов многофакторного SAM-метода каждая временная точка и контрольные данные воспринимались как отдельный класс данных, а для однофакторного временного SAM-метода все данные нормированы с учетом контрольных значений и обрабатывались в совокупности. Выполнена иерархическая кластеризация дифференциально выраженных генов. Оценка качества иерархической кластеризации произведена при помощи кофенетического корреляционного коэффициента и визуально с использованием дендрограмм и графиков разброса в пространстве первых главных компонент. В качестве критерия определения количества кластеров был выбран коэффициент несоответствия (*inconsistency*) [40]. Установлено предпочтение результату, полученному многофакторным SAM-методом (рис. 4, з, д), в силу пространственной разделенности кластеров генов и четкой выраженности областей кластеров. В результате иерархического кластерного анализа определены три главных кластера генов (рис. 5).

Первый кластер (305 генов) характеризует первоначально подавленные гены, экспрессия которых оставалась почти неизменной в течение первых 12 ч. Затем экспрессии этих генов достигают максимума в промежуток времени от 12 до 48 ч после начала воздействия, причем гены становятся выраженными и уже не изменяют свое состояние. Аналогичные результаты наблюдаются при анализе воздействия IFN- γ на второй кластер генов (637 генов), только гены со временем подавляются. Гены, относящиеся к третьему кластеру (149 генов), начинают реагировать на IFN- γ сразу же после начала воздействия и меняют свое состояние с подавленного на выраженное за 12 ч после начала воздействия, после чего с 12 до 24 ч практически не меняют свое состояние, в промежуток времени с 24 до 72 ч – подавлены.

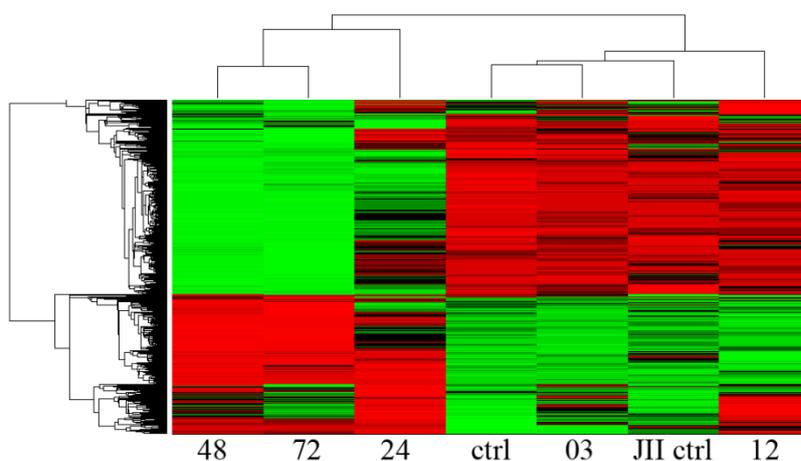
Значения экспрессии генов, которые получены для клеток, обработанных блокирующим сигналом от интерферона ингибитором, аналогичны результатам контрольных необработанных клеток. Это подтверждает вывод о том, что изменения экспрессии обусловлены воздействием IFN- γ на клетки, но не изменениями в самих клетках с течением времени. Данный вывод также подтверждается результатами анализа SAM-методом по каждой временной точке, где количество значимых генов для III ctrl мало. Дендрограмма на рис. 5, в показывает кластеризацию генов не только по изменениям в течение воздействия, но и по временным точкам. Если остановиться на втором уровне разветвления дерева кластеризации, можно выделить четыре основных кластера: первый – 48 и 72 ч; второй – 24 ч; третий – ctrl, 03, III ctrl; четвертый – 12 ч после начала воздействия. На рис. 5, б указан список наиболее известных генов, транскрипционные изменения которых типичны для воздействия клетки IFN- γ и были подтверждены в работах других авторов [19].



a)

Кластеры		
первый	второй	третий
AAK1	A2M	ACSL5
AARS	AASS	ADAP1
ATG12	COL5A2	CXCL10
ATL3	CPVL	CXCL11
DEPDC7	CRISPLD1	IRF1
HDAC9	IFT122	SECTM1
HERPUD1	IGFBP2	SERPING1
HRH1	IGFBP5	SLC12A7
IER2	MMP1	SMOX
IL1A	SOX5	SOCS3
TMCC3	THYN1	SP110
TMEM154	TIAM1	SPRY4
TMEM88	VCAN	STAT1

б)



в)

Рис. 5. Результаты кластеризации: а) дендрограмма кластеризации, показывающая изменения экспрессии генов с течением времени (справа представлены профили экспрессии кластеров); б) список наиболее характерных генов для каждого кластера, названия генов представлены в виде HGNC(HUGO GeneNomenclatureCommittee)-кодов; в) дендрограмма кластеризации профилей экспрессии и временных точек (слева – дерево кластеризации профилей, сверху – дерево кластеризации временных точек)

На завершающем этапе анализа данных определены активные биофункции клетки, контролируемые выраженными генами, полученными двухфакторным парным SAM-методом. Дополнительный анализ выполнен над кластерами генов, полученных многофакторным SAM-методом для определения значимых функций, активность которых изменялась в соответствии с профилями экспрессии кластеров. Статистическая значимость биофункций определялась с помощью точного критерия Фишера ($p < 0,01$). В результате анализа группы генов, полученных двухфакторным парным SAM-методом, определены 2975 значимых функций (2419 биологических и 556 молекулярных), характерных для выраженных генов, и 2706 значимых функций (1487 биологических и 1219 молекулярных функций), характерных для подавленных генов. Часть полученных результатов, которая качественно описывает общую совокупность функций (указаны для каждой группы наиболее значимые функции), можно найти по адресу <http://sablab.net/students/saetchnikov/> в табл. 1. Полученные результаты можно разделить на пять основных групп функций в соответствии с тем, что функции выраженных либо подавленных генов имеют доминирующую значимость в определенный момент времени. В 3 ч после начала воздействия пик выраженности достигают 6,6 % всех статистически значимых функций, в 12 ч – 22,9 %, в 24 ч – 44,9 %, в 48 ч – 14,2 %, в 72 ч – 9,2 %, в 72 ч с ингибитором – 2,2 % всех выраженных функций. Для подавленных функций пик подавленно-

сти в 12 ч после начала воздействия достигают 3,5 % функций, в 24 ч – 19,2 %, в 48 ч – 25,3 %, в 72 ч – 52 % всех выраженных функций.

Абсолютное большинство функций достигают пика подавленности в конце эксперимента, тогда как вплоть до 24 ч после начала эксперимента фактически нет подавленных функций. При этом в промежуток времени с 3 до 24 ч после начала эксперимента происходит наибольший рост количества функций, достигших пика выраженности, а в 72 ч после начала эксперимента пика выраженности достигают лишь 9,2 % всех выраженных функций. В целом наблюдается асимметричный процесс изменения количества значимых функций с течением времени.

В начале процесса воздействия (3–12 ч) выражены биологический процесс реакции на IFN- γ (response to interferon-gamma) и биологический процесс реакции иммунной системы (immune response). Через 12 ч после начала воздействия большинство выраженных функций составляют биологические процессы сигнального пути (signaling pathway). В 24 ч преимущественно выражены молекулярные функции связывания (binding), причем в основном это молекулярные функции связывания белков (protein binding), при этом похожее доминирование можно наблюдать в 72 ч после начала воздействия, но уже для подавленных функций. В 48 ч можно отдельно выделить группу выраженных функций, связанных с положительной регуляцией процессов активности (positive regulation), в 72 ч после начала воздействия наиболее значимы выраженные молекулярные функции связывания ионов (ion binding) и транскрипции (transcription). Для подавленных генов в 12 ч характерны функции, связанные с положительной либо отрицательной регуляцией процессов в клетке (regulation), в 48 ч – связанные с метаболическими процессами в клетке (metabolic process). Для 24 ч после начала воздействия характерны также некоторые функции, связанные с метаболическими процессами, но более значимые, чем функции биосинтетического процесса (biosynthetic process).

Для кластеров получены следующие результаты:

- 1) гены, которые попали в первый кластер, имеют 638 (554 биологических, 84 молекулярных) значимых функций;
- 2) гены, относящиеся ко второму кластеру, имеют 1200 (788 биологических, 412 молекулярных) значимых функций;
- 3) гены третьего кластера имеют 1096 (838 биологических, 258 молекулярных) значимых функций.

В первом кластере большинство активных функций отражают процесс биосинтеза аминокислот (amino acid biosynthetic process). Активность процессов биосинтеза аминокислот увеличивается с 12 ч и достигает максимума в 48 ч после начала воздействия. Во втором кластере можно отдельно выделить функции сборки (assembly) и метаболических процессов (metabolic process). В третьем кластере функции характеризуют иммунную реакцию клетки (innate immune response, immune response, immune response to tumor cell etc) и пути передачи сигнала (interferon-gamma-mediated signaling pathway, tumor necrosis factor-mediated signaling pathway etc). Иммунная реакция клетки постоянно нарастает с начала воздействия и в промежуток времени с 12 до 24 ч с начала воздействия IFN- γ максимально активна. После 24 ч активность иммунной системы постепенно падает, а в 72 ч после начала воздействия иммунная система неактивна.

5. Сравнение результатов

Приведем сравнение полученных и опубликованных в [19] результатов. В силу того что конечные данные (значимые функции) получены с помощью разных методик и в различных программных обеспечениях, напрямую сравнить их затруднительно. Для сравнения количества выраженных функций в заданные моменты времени (03Н, 12Н, 24Н, 48Н, 72Н) по отношению к нулевому отсчету был использован веб-ресурс <http://biocompendium.embl.de/>. Использование данного ресурса позволяет исключить неопределенность, связанную с вопросами стандартизации протокола выделения биофункций в работе [19]. Анализ выполнен как для списков генов, полученных в работе [19], так и для результатов данной работы. Результаты представлены в таблице. Необходимо отметить, что в результате работы

GeneExpressionAnalyser почти по всем временным точкам было получено большее число значимых функций (в среднем на 44 %), чем в работе [19]. Процент общих значимых функций в среднем составляет 68 % от количества значимых функций, которые были получены в результате анализа, проведенного в работе [19].

Количество значимых функций по временным точкам

Источники сравнимых данных	03Н	12Н	24Н	48Н	72Н
Работа [19]	4	16	24	47	34
Результаты данных исследований	3	17	47	66	48
Общие результаты для данных исследований и работы [19]	3	12	15	31	24

Вероятно, большее количество значимых функций, выделенных программным пакетом *GeneExpressionAnalyser*, нежели функций, выделенных с помощью ресурса <http://biocompendium.embl.de/>, является результатом обновления базы данных GO-аннотаций, а также учета молекулярных функций.

Заключение

Разработан программный комплекс *GeneExpressionAnalyser* для анализа широкого набора биочипов ДНК, интегрирующий основные этапы анализа данных, такие как: загрузка данных, предварительная обработка, выделение значимых генов, определение доминирующих функций клетки. Работоспособность алгоритмов программного пакета подтверждена на примерах анализа смоделированных и экспериментальных данных. Пакет *GeneExpressionAnalyser* имеет следующие преимущества над существующим программным обеспечением в данном сегменте экспериментальных исследований:

1) свободно распространяется в пределах Республики Беларусь, требует своевременного обновления базы данных аннотаций. Распространяется программный пакет как исполнительный файл с набором библиотек, что значительно упрощает работу пользователя;

2) позволяет исследовать биочипы Affymetrix, двухцветные биочипы, а также наборы данных, представленные в табличном виде;

3) содержит широкий набор методов многомерного анализа данных;

4) выполняет эффективный анализ аннотаций списка дифференциально выраженных генов.

В работе приведены результаты исследований изменений экспрессии выраженности генов в клетке меланомы A375 под воздействием IFN- γ с течением времени с помощью программного пакета *GeneExpressionAnalyser*. Полученные результаты воспроизводят опубликованные ранее данные, что подтверждает работоспособность разработанных методов анализа данных и программного обеспечения. Дополнительный материал по теме данных исследований можно найти по адресу <http://sablab.net/students/saetchnikov/>.

Список литературы

1. Свешникова, А.Н. Экспрессия генов и микрочипы: проблемы качественного анализа / А. Н. Свешникова, П.С. Иванов // Рос. хим. ж. – 2007. – Т. 51 (1). – С. 127–135.
2. Maciejewski, H. Gene set analysis methods: statistical models and methodological differences / H. Maciejewski // Brief Bioinform. – 2013. – Feb. 14. – P. 1–15.
3. Assessment of gene set analysis methods based on microarray data / H.A. Majd [et al.] // Gene. – 2013. – Vol. 534. – P. 383–389.
4. SplicerEX: a tool for the automated detection and classification of mRNA changes from conventional and splice-sensitive microarray expression data / T.J. Robinson [et al.] // RNA. – 2012. – Vol. 18(8). – P. 1435–1445.
5. BEAT: Bioinformatics Exon Array Tool to store, analyze and visualize Affymetrix GeneChip Human Exon Array data from disease experiments / A. Consiglio [et al.] // BMC Bioinformatics. – 2012 – 13(Suppl 4): S21. – P. 1–14.

6. Bar-Joseph, Z. Studying and modelling dynamic biological processes using time-series gene expression data / Z. Bar-Joseph, A. Gitter, A. Simon // *Nature Reviews Genetics*. – 2012. – Vol. 13. – P. 552–564.
7. Mehta, J.P. Software and tools for microarray data analysis / J.P. Mehta, S. Rani // *Methods Mol Biol*. – 2011. – Vol. 784. – P. 41–53.
8. Pathway analysis software: annotation errors and solutions / N.K. Henderson-Maclennan [et al.] // *Mol Genet Metab*. – 2010. – Vol. 101(2–3) – P. 134–140.
9. Brian, S. E. Handbook of Statistical Analyses Using R / S.E. Brian, T.A. Hothorn. – Chapman and Hall/CRC, 2009. – 376 p.
10. Bioconductor: open software development for computational biology and bioinformatics / R.G. Gentleman [et al.] // *Genome Biology* – 2004. – Feb. 14. – P. 80.1–80.16
11. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor / S. Anders [et al.] // *Nat Protoc*. – 2013. – № 8(9). – P. 1765–1786.
12. BeadArray Expression Analysis Using Bioconductor / M. Dunning [et al.] // *PLoS Comput Biol*. – 2011. – № 7(12). – P. 1–39.
13. Coral: an integrated suite of visualizations for comparing clusterings/ D. Filippova, A. Gadani, C. Kingsford // *BMC Bioinformatics* – 2012. – Vol. 13:276. – P. 1–13.
14. DMET-Analyzer: automatic analysis of Affymetrix DMET Data / P.N. Guzzi [et al.] // *BMC Bioinformatics* – 2012. – Vol. 13:258. – P. 1–10.
15. eXframe: reusable framework for storage, analysis and visualization of genomics experiments/ A.U. Sinha [et al.] // *BMC Bioinformatics* – 2011. – Vol. 12:452. – P. 1–13.
16. Next Generation Sequencing & Microarray Data Analysis Software | Partek Incorporated [Electronic resource]. – 2013. – Mode of access : <http://www.partek.com>. – Date of access : 15.11.2013.
17. Ingenuity IPA – Integrate and understand complex omics data [Electronic resource]. – 2013. – Mode of access : <http://www.ingenuity.com/products/ipa>. – Date of access : 16.11.2013.
18. GoMiner: a resource for biological interpretation of genomic and proteomic data / B.R. Zeeberg [et al.] // *Genome Biol*. – 2003. – Vol. 4 (4). – Art. R28. – P. 1–8.
19. Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function / P.V. Nazarov [et al.] // *Nucleic Acids Research*. – 2013. – Vol. 41(5). – P. 2817–2831.
20. Exploration, Normalization, and Summaries of High Density Oligonucleotide Array Probe Level Data / R.A. Irizarry [et al.] // *Biostatistics*. – 2003. – Vol. 4(2). – P. 249–264.
21. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation / S. Dudout [et al.] // *Nucleic Acids Research*. Oxford University Press. – 2002. – Vol. 30, 4 e15. – P. 1–10.
22. Comparison of Affymetrix data normalization methods using 6,926 experiments across five array generations / R. Autio [et al.] // *BMC Bioinformatics*. – 2009. – 10 (Suppl 1): S24. – P. 1–12.
23. Analysis of boutique arrays: A universal method for the selection of the optimal data normalization procedure / B. Uszczyńska [et al.] // *Mol. Med*. – 2013. – Sep. 32(3). – P. 668–684.
24. Bra's, Lígia P. Improving cluster-based missing value estimation / Lígia P. Brás, José C. Menezes // *Biomolecular Engineering*. – 2007. – T. 24. – P. 273–282.
25. Tusher, V.G. Significance analysis of microarrays applied to the ionizing radiation response / V.G. Tusher, R. Tibshirani, G. Chu // *PNAS*. – 2001. – Vol. 98,9. – P. 5116–5121.
26. Speed, T. Statistical Analysis of Gene Expression Microarray Data: Clustering Microarray-Data / T. Speed // Chapman and Hall/CRC. – 2005. – 240 p.
27. Прикладная статистика: классификация и снижение размерности : справ. изд. / С.А. Айвазян [и др.]. – М. : Финансы и статистика, 1989. – 607 с.
28. The Gene Ontology [Electronic resource]. – 2013. – Mode of access : <http://www.geneontology.org>. – Date of access : 15.11.2013.
29. High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID) / B.R. Zeeberg [et al.] // *BMC Bioinformatics*. – 2005. – Vol. 6; 168. – P. 1–18.

30. Molecular Devices Launches the GenePix(R) SL50 Slide Loader and GenePix(R) Pro 7.2 Software // PR Newswire Europe Including UK Disclose [Electronic resource]. – 2011. – Mode of access : <http://www.prnewswire.co.uk/news-releases/molecular-devices-launches-the-genepixr-sl50-slide-loader-and-genepixr-pro-72-software-145284555.html>. – Date of access : 16.11.2013.
31. Home | Affymetrix [Electronic resource]. – 2013. – Mode of access : <http://www.affymetrix.com>. – Date of access : 13.11.2013.
32. Advanced spot quality analysis in two-colour microarray experiments / M. Yatskou [et al.] // BMC Research Notes. – 2008. – Vol. 1:80. – P. 1–13.
33. Dynamic regulation of microRNA expression following Interferon- γ -induced gene transcription / S. Reinsbach [et al.] // RNA Biology. – 2012. – Vol. 9:7. – P. 978–989.
34. Samr: SAM: Significance Analysis of Microarrays [Electronic resource]. – 2011. – Mode of access : <http://cran.r-project.org/web/packages/samr/samr.pdf>. – Date of access : 13.11.2013.
35. Data Mining Practical Machine Learning Tools and Techniques / Ian H. Witten [et al.]. – The Morgan Kaufmann Series in Data Management Systems, 2011. – 664 p.
36. Hyvarinen, F. Independent Component Analysis / F. Hyvarinen. – Wiley series, 2001. – 505 p.
37. Разработка метода главных компонент для анализа микрочипов ДНК / А.В. Саечников // 69-я научная конф. студентов и аспирантов БГУ : тез. докл. – Минск, 2012. – С. 268–272.
38. Саечников, А.В. Программный пакет *GeneExpressionAnalyser* для анализа микрочипов ДНК / А.В. Саечников, Н.Н. Яцков, В.В. Апанасович // Медэлектроника 2012 : тез. докл. – Минск, 2012. – С. 79–81.
39. Novikov, E. An algorithm for automatic evaluation of the spot quality in two-color DNA microarray experiments / E. Novikov, E. Barillot // BMC Bioinformatics. – 2005. – Vol. 6: 293. – P. 1–18.
40. Uragun, B. The discrimination of interaural level difference sensitivity functions: development of a taxonomic data template for modelling / B. Uragun, R. Rajan // BMC Neuroscience. – 2013. – Vol. 14: 144. – P. 1–19.

Поступила 26.12.2013

¹*Белорусский государственный университет,
Минск, ул. Курчатова, д. 1
e-mail: saetchnikov.anton@tut.by*

²*Центр геномных исследований,
L-1526 Люксембург
e-mail: petr.nazarov@crp-sante.lu*

A.V. Saetchnikov, M.M. Yatskou, P.V. Nazarov, L. Vallar, V.V. Apanasovich

ANALYSIS OF CELLULAR REACTION TO IFN- γ STIMULATION BY A SOFTWARE PACKAGE *GeneExpressionAnalyser*

The software package *GeneExpressionAnalyser* for analysis of the DNA microarray experimental data has been developed. The algorithms of data analysis, differentially expressed genes and biological functions of the cell are described. The efficiency of the developed package is tested on the published experimental data devoted to the time-course research of the changes in the human cell under the influence of IFN- γ on melanoma. The developed software has a number of advantages over the existing software: it is free, has a simple and intuitive graphical interface, allows to analyze different types of DNA microarrays, contains a set of methods for complete data analysis and performs effective gene annotation for a selected list of genes.

УДК 681.382

А.Г. Стрижнев, Г.В. Ледник, А.А. Шихов, А.Н. Русакович

КОМПЕНСАЦИЯ ОШИБКИ ПЕРЕХОДА ЦИФРОВОЙ СЛЕДЯЩЕЙ СИСТЕМЫ ПРИ КРУГОВОМ СЛЕЖЕНИИ

Рассматривается алгоритм компенсации ошибки перехода, которая возникает при работе цифровой следящей системы в режиме кругового слежения за периодическим сигналом. Особенность данного алгоритма заключается в простой реализации при сохранении высокой устойчивости работы в режиме кругового слежения и наличии возмущений колебательного характера. Дается сравнительная оценка полученного алгоритма и существующих решений. Путем математического моделирования осуществляется проверка эффективности разработанного алгоритма. Предлагается к практическому применению структурная схема цифровой следящей системы, реализующая данный алгоритм компенсации.

Введение

Развитие средств вычислительной техники, особенно микропроцессоров и микроЭВМ, послужило основой широкого распространения цифровых систем автоматического управления (САУ). Описание моделей объектов управления и корректирующих устройств в дискретном времени позволяет существенно упростить синтез регуляторов и их техническую реализацию. Многие алгоритмы функционирования регуляторов, применяемые при расчете цифровых систем, часто могут быть реализованы только на микроконтроллерах или микроЭВМ. На основе цифрового управления могут быть построены САУ любых типов, а программное обеспечение систем можно корректировать как при проектировании, так и в процессе эксплуатации [1].

Одним из видов цифровых САУ являются цифровые следящие системы (ЦСС), которые в режиме кругового слежения широко используются на радиолокационных станциях, в оптико-механических системах обзора и другой технике [2, 3]. При наличии внешних возмущений колебательного характера (например, порывов ветра и др.) из-за инерционности механических элементов возникает ошибка перехода, которая снижает устойчивость слежения, а иногда делает невозможным круговое синхронное слежение. Существуют технические решения [4, 5], позволяющие повысить точность и устойчивость работы ЦСС при работе в режиме кругового слежения, однако данные решения либо не обеспечивают устойчивой работы ЦСС при наличии возмущений колебательного характера, либо сложны в реализации.

В настоящей работе проводится анализ эффективности применения существующих технических решений и предлагается новый алгоритм компенсации ошибки перехода.

1. Цифровая следящая система

ЦСС обычно имеет структурную схему [6, 7], которая представлена на рис. 1. Система содержит устройство управления УУ, цифровое вычислительное устройство ЦВУ, которое включает в себя цифровой вычислитель ЦВ и цифровой регулятор ЦР, усилительно-преобразовательное устройство УПУ, исполнительный двигатель ИД, приводной механизм ПМ, цифровой преобразователь угла ЦПУ и нагрузку Н.

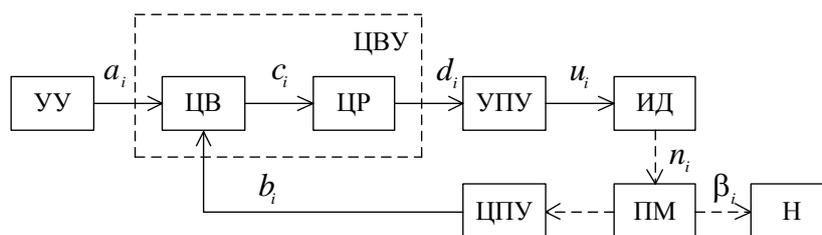


Рис. 1. Типовая структурная схема ЦСС

Система работает следующим образом. Устройство УУ вырабатывает управляющий сигнал a_i в виде цифрового кода, который поступает на суммирующий вход устройства ЦВ, где из управляющего сигнала вычитается сигнал b_i обратной связи, поступающий из преобразователя ЦПУ. Сигнал разности c_i с выхода устройства ЦВ, вычисленный по формуле $c_i = a_i - b_i$, поступает на регулятор ЦР, который обеспечивает требуемые динамические показатели системы и формирует сигналы в виде цифрового кода d_i для управления работой устройства УПУ. Устройство УПУ усиливает и преобразует эти сигналы в напряжения u_i , необходимые для управления двигателем ИД. Двигатель ИД через механизм ПМ поворачивает нагрузку и входную ось преобразователя ЦПУ на такой угол β_i , при котором $c_i = 0$. Устройства УПУ, ИД, ПМ, Н и ЦПУ образуют объект управления (ОУ) системы.

В установившемся режиме работы системы (рис. 1) предполагается, что сигнал c_i (ошибка системы) будет минимальным и не превысит заданного значения. Однако в реальности это не так. При работе системы в режиме кругового слежения за периодическим сигналом из-за инерционности элементов (двигателя ИД, механизма ПМ, нагрузки Н) имеет место временное запаздывание сигнала b_i обратной связи от управляющего сигнала a_i . Это приводит к тому, что при переходе значений цифровых кодов a_i и b_i с максимальных на минимальные или наоборот возникает ошибка перехода, которая значительно превышает установившуюся в линейной зоне динамическую ошибку c_i системы и имеет противоположный знак.

Временные диаграммы, поясняющие работу системы (рис. 1) в режиме кругового слежения, показаны на рис. 2.

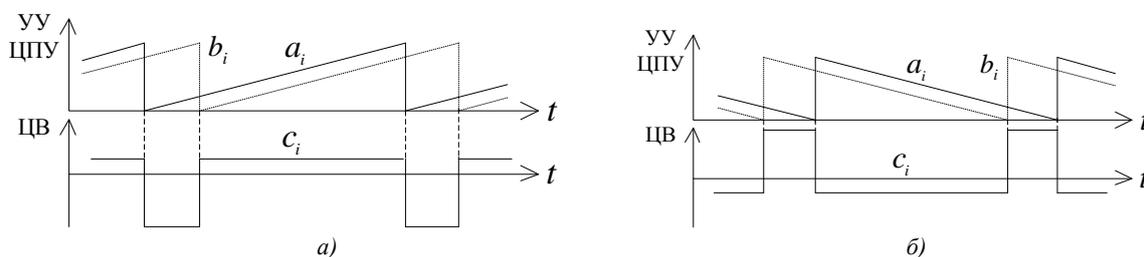


Рис. 2. Временные диаграммы для режима кругового слежения и различных направлений сигналов: а) прямого; б) обратного

Наличие данного недостатка приводит к появлению рывков при круговом слежении. В связи с этим возникает необходимость компенсации ошибки перехода и обеспечения устойчивой работы системы в режиме кругового слежения. Наиболее близкие существующие [5, 6] способы компенсации ошибки перехода предполагают изменение структурной схемы ЦСС. Эти способы отличаются не только структурной схемой, но и принципом работы, поэтому возникает необходимость их подробного рассмотрения.

2. Компенсация ошибки перехода путем запоминания текущей ошибки

Структурная схема ЦСС [4], реализующая компенсацию ошибки перехода путем запоминания текущей ошибки (рис. 3), в отличие от системы (рис. 1) дополнительно содержит цифровой регистр памяти РП, триггер Т, четыре элемента И1 – И4 совпадения, два логических элемента 2-ИИ-ИЛИ ЛЭ1 и ЛЭ2, два формирователя ФИ1 и ФИ2 импульсов по переднему фронту и срезу соответственно.

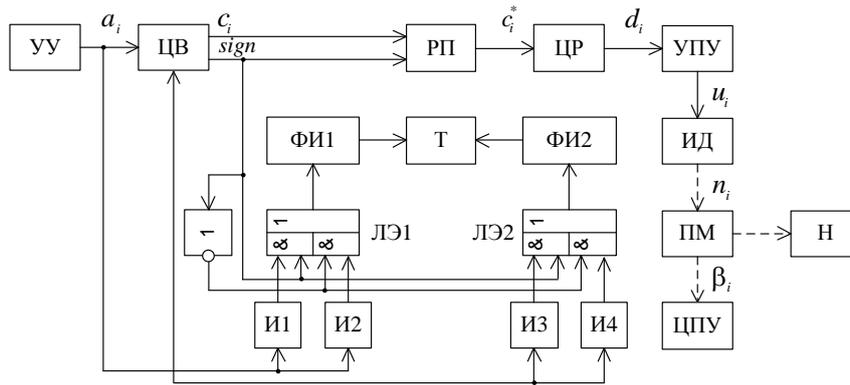


Рис. 3. Структурная схема ЦСС с компенсацией ошибки перехода путем запоминания текущей ошибки

Работа системы (рис. 3) на линейном участке ничем не отличается от работы рассмотренной ранее системы (рис. 1).

В режиме кругового вращения управляющий сигнал a_i и сигнал b_i обратной связи поступают на входы элементов совпадения И1, И2 и И3, И4 соответственно. Выделенные элементами совпадения четыре сигнала (элементом И1 – a_{\max} , элементом И2 – a_{\min} , элементом И3 – b_{\min} , элементом И4 – b_{\max}) поступают на входы логических элементов ЛЭ1 и ЛЭ2.

При прямом изменении значений управляющего сигнала a_i (... , $a_{\max} - 1$, a_{\max} , a_{\min} , $a_{\min} + 1$, ...) в момент совпадения сигнала «знак разности» с сигналом a_{\max} на выходе элемента ЛЭ1 выделяется сигнал, который после прохождения формирователя ФИ1 импульсов по фронту устанавливает триггер Т в состояние логической единицы. В момент совпадения сигнала «знак разности» с сигналом b_{\min} сигнал, выделенный на выходе элемента ЛЭ2, после прохождения формирователя ФИ2 импульсов по срезу устанавливает триггер Т в состояние логического нуля (рис. 4, а).

При обратном изменении управляющего сигнала a_i (... , $a_{\min} + 1$, a_{\min} , a_{\max} , $a_{\max} - 1$, ...) в момент совпадения инверсного сигнала «знак разности» с сигналом a_{\min} на выходе элемента ЛЭ1 также выделяется сигнал, который устанавливает триггер Т в состояние логической единицы. В состояние логического нуля триггер Т возвращается при совпадении инверсного сигнала «знак разности» с сигналом b_{\max} (рис. 4, б).

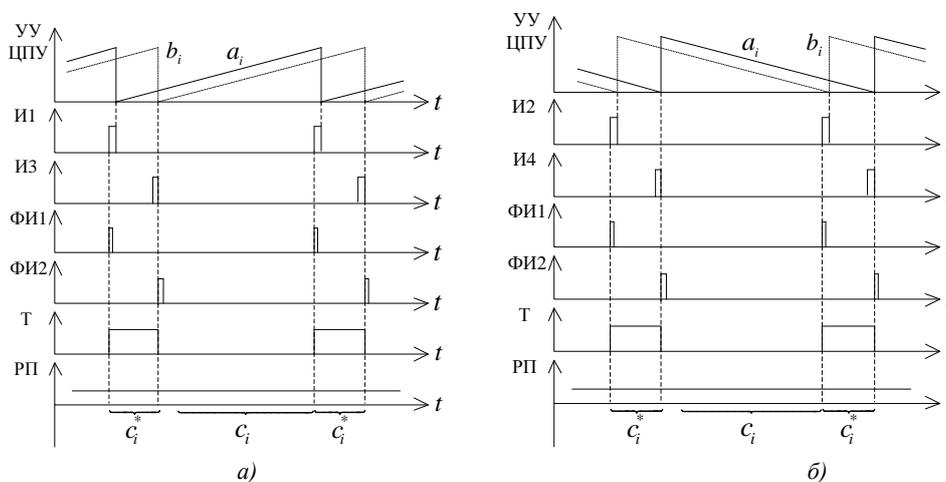


Рис. 4. Временные диаграммы работы ЦСС с компенсацией ошибки перехода путем запоминания текущей ошибки для различных направлений кругового слежения: а) прямого; б) обратного

Во время действия импульса, сформированного триггером Т, регистр памяти РП запоминает значение ошибки c_i системы и сигнал «знак разности», которые имели место при a_{\max} (для прямого направления вращения) или при a_{\min} (для обратного направления вращения). В результате ошибка перехода не поступает на регулятор ЦР и прямой контур системы обрабатывает запомненную ранее текущую ошибку. Благодаря этому повышается устойчивость работы системы при переходе числовых последовательностей сигналов задающего устройства и преобразователя угол-код с максимального на минимальное значение или наоборот.

Следует заметить, что система (рис. 3) хорошо работает при непрерывном круговом синхронном слежении с постоянной скоростью и отсутствии возмущений, имеющих колебательный характер. В ходе экспериментальных исследований было выявлено, что при наличии колебательных возмущений в момент обработки системой запомненной ошибки могут произойти изменения направлений слежения и сбой работы системы. Значительно лучшей устойчивостью работы в режиме кругового слежения и при наличии колебательных воздействий обладает система, в которой компенсация ошибки перехода осуществляется путем выделения сигнала перехода [5].

3. Компенсация ошибки перехода путем выделения сигнала перехода

Структурная схема ЦСС [5], реализующая компенсацию ошибки перехода путем выделения сигнала перехода (рис. 5), в отличие от системы (рис. 1) дополнительно содержит два цифровых сумматора ЦС1 и ЦС2, четыре цифровых компаратора ЦК1 – ЦК4, два элемента совпадения И1 и И2.

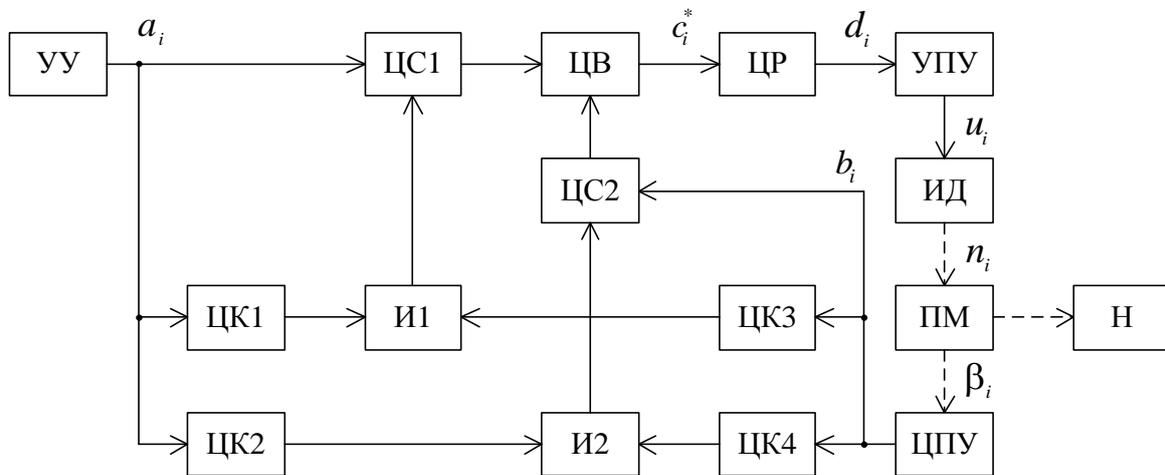


Рис. 5. Структурная схема ЦСС с компенсацией ошибки перехода путем выделения сигнала перехода

Временные диаграммы (рис. 6) поясняют работу структурной схемы ЦСС. Цифровой компаратор (рис. 7) содержит два элемента совпадения И3 и И4 и триггер Т.

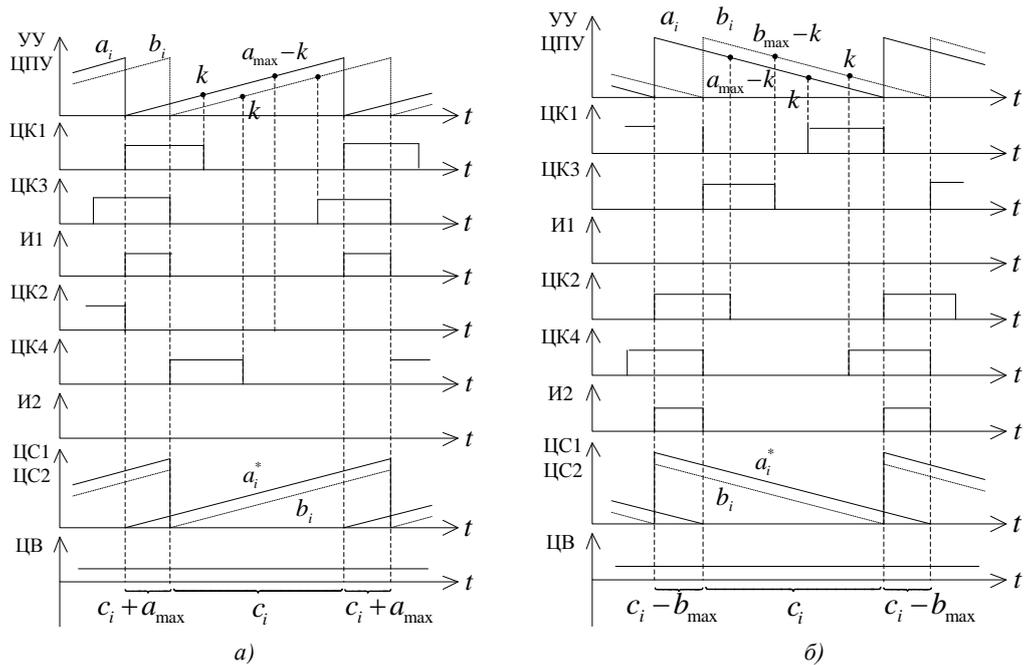


Рис. 6. Временные диаграммы работы ЦСС с компенсацией ошибки перехода путем выделения сигнала перехода для различных направлений кругового слежения: а) прямого; б) обратного

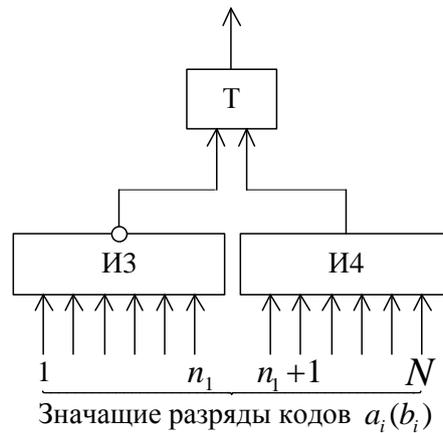


Рис. 7. Структурная схема цифрового компаратора

Работа системы (рис. 5) на линейном участке ничем не отличается от работы рассмотренной ранее системы (рис. 1).

Во время работы системы (рис. 5) в режиме кругового слежения управляющий сигнал a_i и сигнал b_i обратной связи поступают на входы цифровых компараторов ЦК1, ЦК2 и ЦК3, ЦК4 соответственно. Цифровые компараторы ЦК1 и ЦК4 вырабатывают логическую единицу в диапазоне $[\min; k]$, а цифровые компараторы ЦК2 и ЦК3 – в диапазоне $[\max-k; \max]$ значений кодов a_i и b_i соответственно, где k – граница срабатывания цифровых компараторов, которая выбирается в диапазоне $\min - 0,5\max$ (оптимальное значение $k = 1/3\max$) значений кодов a_i или b_i .

За счет инерционности элементов (исполнительного двигателя, приводного механизма, нагрузки и др.) имеет место запаздывание во времени выходного сигнала b_i относительно управляющего сигнала a_i . Это приводит к тому, что импульсы, выделенные цифро-

выми компараторами ЦК1, ЦК2 и ЦК3, ЦК4, в определенные моменты будут совпадать во времени.

При прямом изменении значений управляющего сигнала a_i ($\dots, a_{\max} - 1, a_{\max}, a_{\min}, a_{\min} + 1, \dots$) совпадут сигналы, снимаемые с выходов цифровых компараторов ЦК1 и ЦК3, а значит, элемент И1 сработает и выдаст сигнал перехода в виде логической единицы на второй вход сумматора ЦС1 (см. рис. 6, а).

При обратном изменении управляющего сигнала a_i ($\dots, a_{\min} + 1, a_{\min}, a_{\max}, a_{\max} - 1, \dots$) совпадут сигналы, снимаемые с выходов цифровых компараторов ЦК2 и ЦК4, а значит, элемент И2 сработает и выдаст сигнал перехода в виде логической единицы на второй вход сумматора ЦС2 (см. рис. 6, б).

Наличие на выходах элементов И1 или И2 сигнала перехода в виде логической единицы означает подачу максимальных значений кодов a_i или b_i на вторые входы цифровых сумматоров ЦС1 или ЦС2 соответственно.

В результате цифровые сумматоры ЦС1 или ЦС2 будут производить сложение $a_i + a_{\max}$ ($b_i + b_{\max}$) при наличии сигнала перехода или $a_i + a_{\min}$ ($b_i + b_{\min}$) при его отсутствии.

Таким образом, несмотря на наличие в числовых последовательностях входного сигнала a_i и сигнала b_i обратной связи перехода с максимального на минимальное значение или наоборот, сигнал ошибки, снимаемый с выхода цифрового вычитателя ЦВ, будет соответствовать реальному. Это значит, что ЦСС в режиме кругового вращения и при наличии колебательных возмущений будет работать устойчиво в широком диапазоне скоростей и без рывков.

Следует заметить, что система (рис. 5) в отличие от системы (рис. 3) значительно лучше работает при непрерывном круговом синхронном слежении с различной скоростью и при наличии возмущений, имеющих колебательный характер. Основным недостатком системы (рис. 5) является ее сложность. Структура системы может быть существенно упрощена, если для компенсации ошибки перехода использовать сигнал ошибки системы.

4. Компенсация ошибки перехода с использованием сигнала ошибки системы

С целью упрощения системы (рис. 5) при сохранении высокой устойчивости работы в режиме кругового слежения и при наличии возмущений, имеющих колебательный характер, была разработана ЦСС, реализующая компенсацию ошибки перехода с использованием сигнала ошибки системы (рис. 8). Временные диаграммы, поясняющие ее работу, показаны на рис. 9.

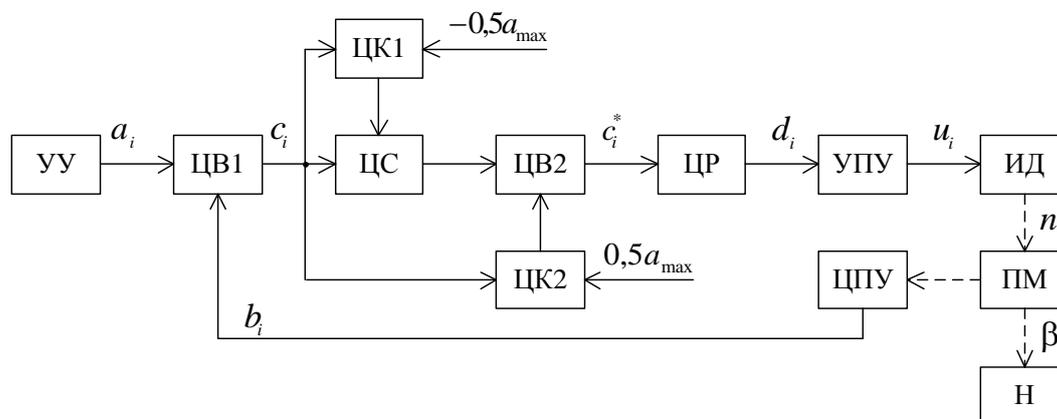


Рис. 8. Структурная схема ЦСС, компенсирующая ошибку перехода с использованием сигнала ошибки системы

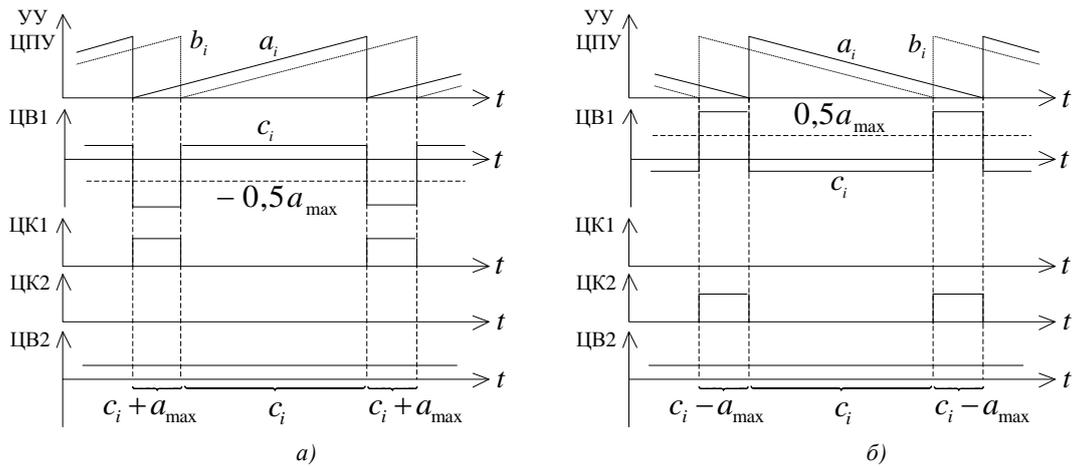


Рис. 9. Временные диаграммы работы ЦСС, компенсирующей ошибку перехода с использования сигнала ошибки системы для различных направлений кругового слежения: а) прямого; б) обратного

Система (рис. 8) в отличие от системы (рис. 1) дополнительно содержит цифровой сумматор ЦС, цифровой вычитатель ЦВ2 и два цифровых компаратора ЦК1 и ЦК2. Цифровой компаратор имеет два цифровых входа (информационный и опорный) и один логический выход.

Работа системы (рис. 8) на линейном участке ничем не отличается от работы рассмотренной ранее системы (рис. 1).

Во время работы системы (рис. 8) в режиме кругового слежения за счет инерционности элементов (двигателя ИД, механизма ПМ, нагрузки Н и др.) имеет место запаздывание во времени выходного сигнала b_i относительно управляющего сигнала a_i . Это приводит к тому, что при переходе числовых последовательностей устройства УУ и преобразователя ЦПУ с максимального на минимальное значение или наоборот на выходе вычитателя ЦВ возникает сигнал ошибки $|c_i| > 0,5a_{\max}$ и компараторы ЦК1 или ЦК2 будут срабатывать.

При прямом изменении значений управляющего сигнала a_i ($\dots, a_{\max} - 1, a_{\max}, a_{\min}, a_{\min} + 1, \dots$) в работе участвует компаратор ЦК1, который сравнивает текущий сигнал ошибки c_i , действующий на информационном входе, с постоянным установочным сигналом $-0,5a_{\max}$, действующим на опорном входе. Когда сигнал ошибки $c_i < -0,5a_{\max}$, компаратор ЦК1 сработает и выдаст логическую единицу на второй вход сумматора ЦС (рис. 9, а).

При обратном изменении значений управляющего сигнала a_i ($\dots, a_{\min} + 1, a_{\min}, a_{\max}, a_{\max} - 1, \dots$) в работе участвует компаратор ЦК2, который сравнивает текущий сигнал ошибки c_i , действующий на информационном входе, с постоянным установочным сигналом $0,5a_{\max}$, действующим на опорном входе. Когда сигнал ошибки $c_i > 0,5a_{\max}$, цифровой компаратор ЦК2 сработает и выдаст логическую единицу на вычитающий вход второго вычитателя ЦВ2 (рис. 9, б).

Наличие на выходе компараторов ЦК1 или ЦК2 логической единицы означает подачу кода a_{\max} , а отсутствие – подачу кода a_{\min} на второй вход сумматора ЦС или вычитающий вход второго вычитателя ЦВ2 соответственно. В результате цифровой сумматор ЦС будет выполнять сложение $c_i + a_{\max}$, а цифровой вычитатель ЦВ2 – вычитание $c_i - a_{\max}$ при наличии логической единицы или $c_i + a_{\min}$, $c_i - a_{\min}$ при ее отсутствии соответственно.

В результате ошибка перехода не возникает и прямой контур системы обрабатывает обычную текущую ошибку. Благодаря этому обеспечивается устойчивая работа системы при переходе числовых последовательностей сигналов задающего устройства и преобразователя угол-код с максимального на минимальное значение или наоборот.

Устойчивая работа системы полностью сохраняется при запаздывании значений сигнала b_i обратной связи от значений управляющего сигнала a_i , а также при наличии колебательных возмущений в режиме кругового слежения. Независимо от того какая возникает по величине и знаку ошибка в процессе работы, система всегда обрабатывает координату в сторону, меньшую $0,5a_{\max}$. Структурная схема системы (рис. 8) значительно проще предыдущей схемы (рис. 5). Для качественной оценки разработанного алгоритма компенсации ошибки перехода необходимо провести проверку системы (рис. 8) в различных условиях работы, которую проще всего осуществить с помощью имитационного моделирования.

5. Имитационное моделирование

Моделирование работы системы проведено в среде Simulink [8] пакета программ MATLAB на примере объекта управления, математическая модель которого описывается передаточной функцией

$$G(s) = \alpha[s(s + a)(s + b)]^{-1}, \tag{1}$$

где $\alpha = 4971,15 \text{ c}^{-3}$, $a = 48,54 \text{ c}^{-1}$, $b = 1,415 \text{ c}^{-1}$.

ОУ имеет цифровые вход и выход с квантованием по времени 0,001 с и уровню 1 дел. и обладает нелинейностью типа «насыщение», которая составляет ± 255 дел., что соответствует максимальному напряжению питания ИД. Коррекция динамических свойств системы осуществляется с помощью оптимального цифрового регулятора [1] с передаточной функцией

$$W(z) = K_0 \frac{1 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}, \tag{2}$$

где $K_0 = 39$; $a_1 = 0,7908$; $a_2 = 0,1271$; $b_1 = -1,3509$; $b_2 = 0,3682$; $h = 0,02 \text{ c}$.

На вход системы подается периодический сигнал a_i , изменяющийся в диапазоне $0 \dots a_{\max}$, где $a_{\max} = 2^{16}$ дел. Максимальное значение сигнала a_i соответствует углу 360° . Для проверки были выбраны две скорости кругового слежения $\omega_1 = 728 \text{ дел./с}$ и $\omega_2 = 7280 \text{ дел./с}$.

Схема моделирования системы с ОУ $G(s)$ и цифровым регулятором $W(z)$ показана на рис. 10, а схемы блоков системы – на рис. 11.

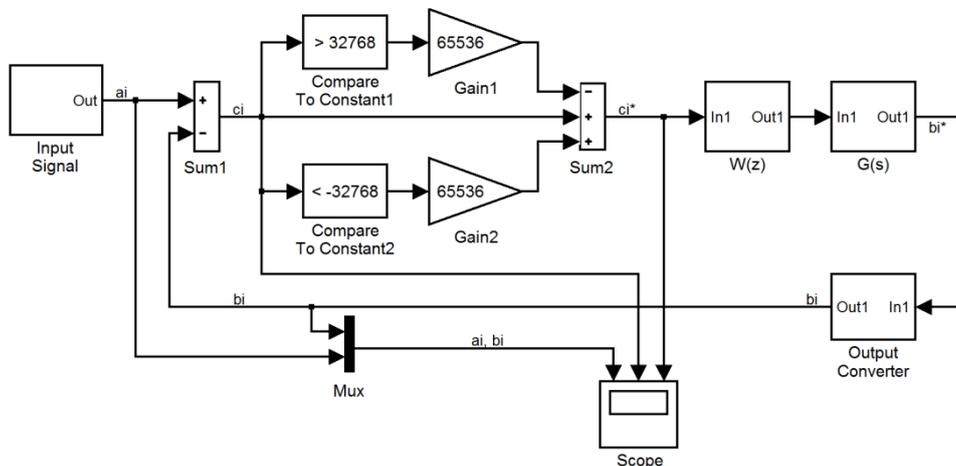


Рис. 10. Схема моделирования работы системы

На схеме моделирования (рис.10), выполненной согласно схеме (рис. 8), устройство УУ представлено блоком Input Signal, устройство ЦВ1 – блоком Sum1, регулятор ЦР – блоком $W(z)$, ОУ – блоком $G(s)$, преобразователь ЦПУ – блоком Output Converter. Устройства, обеспечивающие компенсацию ошибки перехода (ЦК1, ЦК2, ЦС, ЦВ2), реализованы с помощью блоков Compare To Constant1, 2 с параметрами $>2^{15}$ и $<-2^{15}$ соответственно, усилителей Gain1, 2 с коэффициентом усиления 2^{16} и сумматора (вычитателя) Sum2.

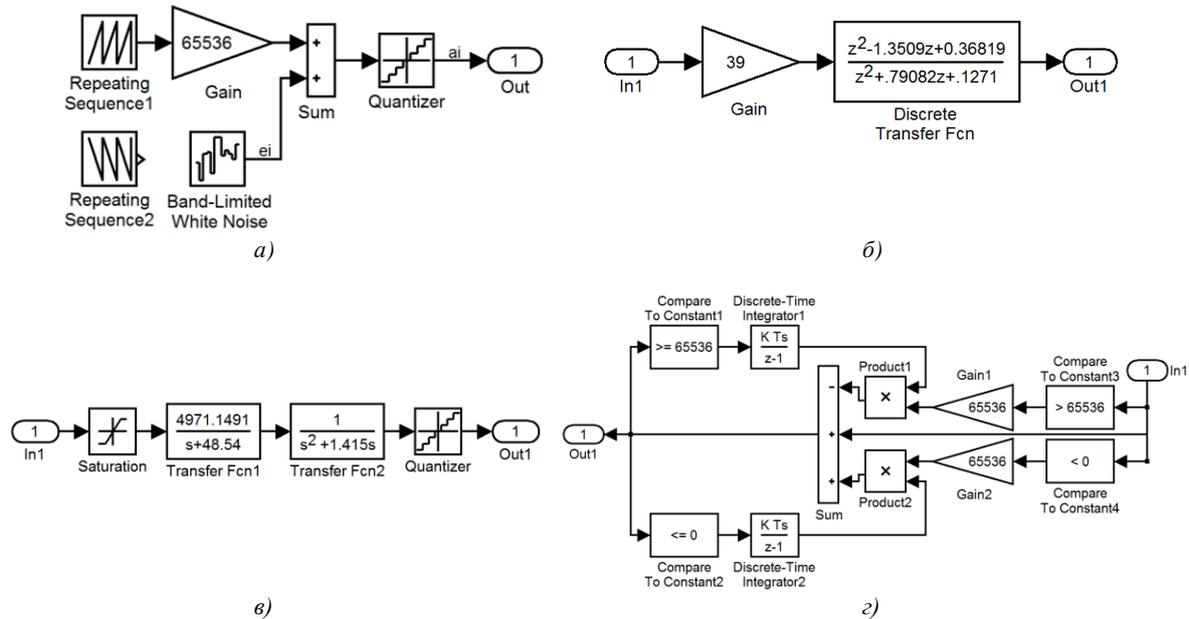


Рис. 11. Модели составных блоков схемы моделирования работы системы: а) УУ; б) ЦР; в) ОУ; г) ЦПУ

Устройство УУ (рис.11, а) состоит из блоков Repeating Sequence 1, 2 с параметрами Time Values = [0 9] для ω_1 и [0 90] для ω_2 , Output Values = [0 1] для прямого и [1 0] для обратного направлений кругового слежения; усилителя Gain с коэффициентом усиления 2^{16} ; сумматора Sum; квантователя по уровню и времени Quantizer с параметрами Quantization interval = 1 и Sample time = 0,001. Для имитации колебательного возмущения к задающему сигналу добавляется сигнал e_i вида «дискретный белый шум», который реализован с помощью блока Band-Limited White Noise с параметрами Noise Power = 0,5, Sample time = 0,001, Seed = 23 341.

Регулятор ЦР (рис.11, б) представлен блоками Discrete Transfer Function и Gain с параметрами, приведенными в (2).

ОУ (рис.11, в) содержит нелинейность типа «насыщение», блок Saturation с параметрами Upper limit = 255 и Lower Limit = -255, блоки Transfer Fcn1, Fcn2 с параметрами, приведенными в (1), блок оцифровки Quantizer с параметрами Quantization interval = 1 и Sample time = 0,001.

Преобразователь ЦПУ (рис.11, г) реализован в виде однооборотного абсолютного энкодера с разрешением 16 бит и содержит следующие блоки: компараторы Compare To Constant1...4 с параметрами $\geq 2^{16}$, ≤ 0 , $> -2^{16}$ и < 0 соответственно; цифровые интеграторы Discrete-Time Integrator1, 2 с параметрами Gain value = 1000 и Sample time = 0,001; умножители Product1, 2; сумматор Sum и усилители Gain1, 2 с коэффициентом 2^{16} . Максимальное значение выходного сигнала ЦПУ соответствует углу 360° .

Результаты моделирования для прямого и обратного направлений кругового слежения со скоростями ω_1 и ω_2 показаны на рис. 12. По оси абсцисс отложено текущее время в секундах, по оси ординат – текущие значения сигналов управления a_i , выходного b_i , ошибок c_i и c_i^* в делениях.

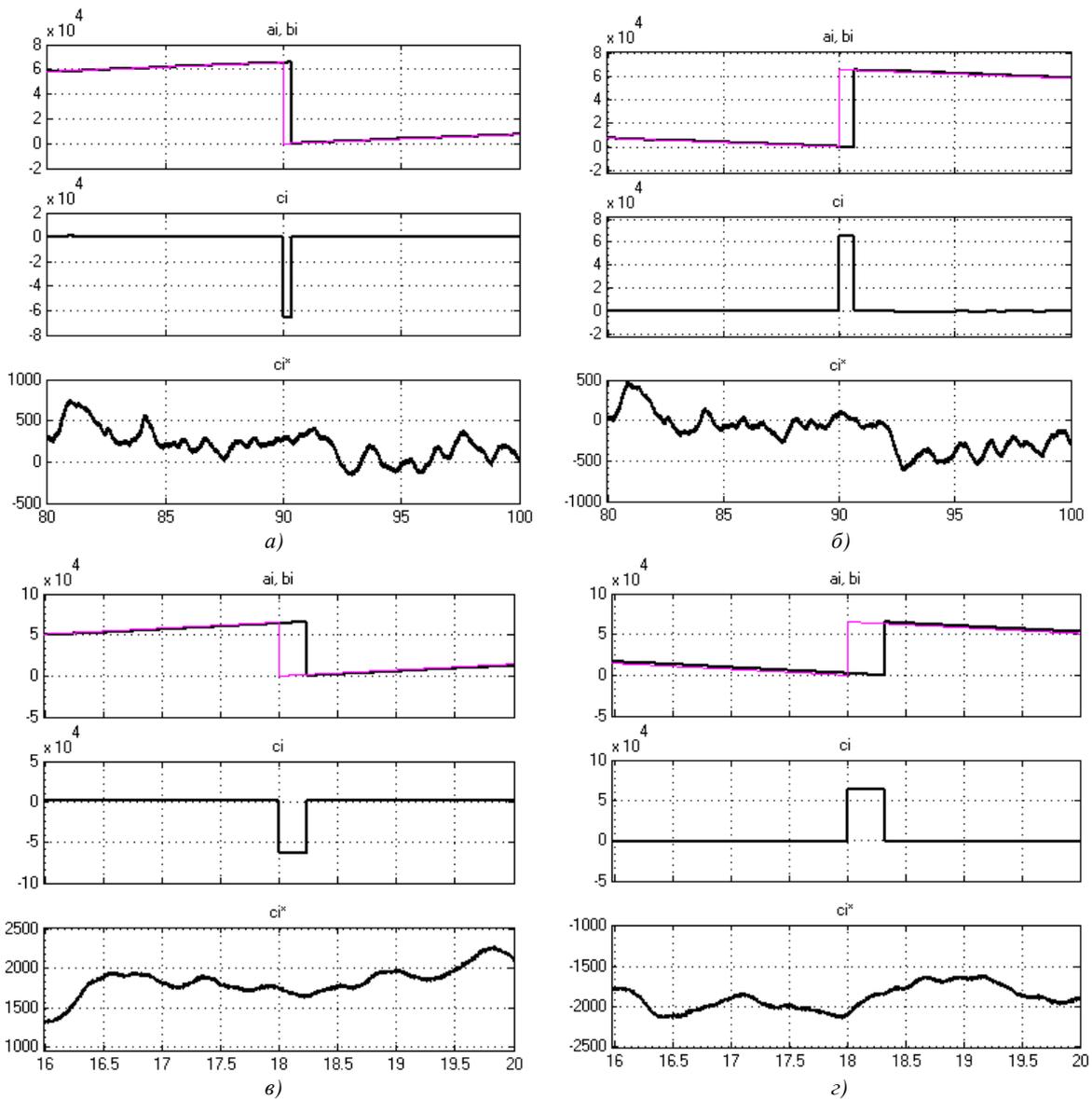


Рис. 12. Результаты моделирования работы системы:

- а) прямое направление движения, скорость ω_1 ; б) обратное направление движения, скорость ω_1 ;
- в) прямое направление движения, скорость ω_2 ; з) обратное направление движения, скорость ω_2

В процессе моделирования получены следующие результаты. При прямом и обратном изменениях задающего сигнала a_i динамическая ошибка в линейной зоне кругового слежения (рис. 12, а, б) со скоростью ω_1 составляет величину $|c_i| \leq 750$ дел., со скоростью ω_2 (рис. 12, в, з) – величину $|c_i| \leq 2300$ дел. Следует заметить, что при отсутствии шума величина динамической ошибки системы значительно меньше и составляет $|c_i| \leq 29$ дел. для скорости ω_1 и $|c_i| \leq 290$ дел. для скорости ω_2 . При переходе сигнала a_i с максимального на минимальное значение или наоборот на выходе блока Sum1 возникает значительный всплеск ошибки в системе до значения $c_i \approx 2^{16}$ дел. Несмотря на это, скомпенсированная ошибка c_i^* на выходе блока Sum2 не изменяется и остается в тех же пределах, что и в зоне линейного слежения. Результаты моделирования подтверждают работоспособность и устойчивость работы системы с

данным алгоритмом компенсации ошибки перехода, а также правильность положений и выводов, сделанных ранее.

Заключение

Проведенный анализ работы существующих технических решений помог найти новый алгоритм компенсации ошибки перехода, основанный на использовании сигнала ошибки системы. Данный алгоритм реализован в виде структурной схемы ЦСС, которая отличается от рассмотренных прототипов значительной простотой и устойчивостью работы. Работоспособность и устойчивость системы с данным алгоритмом компенсации подтверждены результатами имитационного моделирования. Слаботочная часть предложенной ЦСС может быть реализована аппаратно или программно с использованием микроконтроллерной и микропроцессорной техники. Данное техническое решение рекомендуется к широкому практическому применению в технике и промышленности.

Список литературы

1. Гостев, В.И. Системы автоматического управления с цифровыми регуляторами : справ / В.И. Гостев, В.К. Стеклов. – Киев : Радиоаматор, 1998. – 704 с.
2. Ллойд, Дж. Системы тепловидения / Дж. Ллойд. – М. : Мир, 1978. – 407 с.
3. Охрименко, А.Е. Основы радиолокации и радиоэлектронная борьба / А.Е. Охрименко. – М. : Военное издательство министерства обороны СССР, 1983. – 457 с.
4. Цифровая следящая система : а. с. 1149216 СССР, МПК G 05 B 15/02 / В.М. Богуш, А.Г. Стрижнев; ВНИИПИ. – № 3607732/24-24; заявл. 08.04.83; опублик. 07.04.85 // Бюл. № 13. – 1985. – 3 с.
5. Цифровая следящая система : а. с. 1200716 СССР, МПК G 05 B 11/01 / А.Г. Стрижнев; ВНИИПИ. – № 3723269/24-24; заявл. 02.02.84; опублик. 22.08.85. – 9 с.
6. Fadali, M.S. Digital Control Engineering: Analysis and Design / M.S. Fadali, A. Visioli. – Burlington : Elseiver, 2009. – 536 p.
7. Цифровые электроприводы с транзисторными преобразователями / С.Г. Герман-Галкин [и др.]. – Л. : Энергоатомиздат. Ленингр. отд-ние, 1986. – 248 с.
8. Черных, И.В. SIMULINK: среда создания инженерных приложений / И.В. Черных ; под общ. ред. к.т.н. В.Г. Потемкина. – М. : ДИАЛОГ–МИФИ, 2003. – 496 с.

Поступила 4.09.2013

НПООО «ОКБ Техносоюзпроект»,
Минск, пр. Независимости, 115
e-mail: aliaksei.rusakovich@gmail.com.

A.G. Stryzhniou, G.V. Lednik, A.A. Shykhau, A.N. Rusakovich

COMPENSATION OF THE DIGITAL SERVOSYSTEM CIRCULAR TRACKING TRANSITION ERROR

The paper presents a new algorithm for digital servosystem circular tracking transition error compensation. This algorithm is easily realized and ensures high system stability in the circular mode of operation with oscillating perturbations. Comparative assessment of this algorithm with other solving procedures for this problem is made. Digital modeling is applied to verify the effectiveness of this algorithm. A block diagram that realizes this algorithm is presented to be used in a practical application.

УДК 629.33.021:004.94

С.В. Хитриков

МОБИЛЬНАЯ ИЗМЕРИТЕЛЬНАЯ МНОГОКАНАЛЬНАЯ СИСТЕМА ДЛЯ ПРОВЕДЕНИЯ СТЕНДОВЫХ И НАТУРНЫХ ИСПЫТАНИЙ

Описывается измерительная система, предназначенная для проведения натурных испытаний узлов и агрегатов мобильных машин с целью проверки соответствия разработанного изделия нормам технического задания. Система спроектирована и изготовлена на современной элементной базе, которая позволила придать ей уникальные свойства: многоканальность, универсальность, синхронность и удаленное управление.

Введение

Использование компьютерных технологий в задачах проектирования автомобилей вызвало необходимость разработки электронных измерительных и управляющих комплексов для проведения стендовых и дорожных испытаний [1]. Представленная вниманию читателей работа посвящена разработке таких комплексов на основе современных достижений в области электроники. Прототипом данной системы является автоматизированная система виброиспытаний [2], разработанная в Институте технической кибернетики АН БССР с использованием стандарта КАМАК [3, 4].

Фирмы-производители современного измерительного оборудования предлагают большой выбор плат ввода-вывода и нормализаторов сигналов. Разнообразие форм-факторов и компоновки плат может поставить в тупик даже самых профессиональных разработчиков измерительного оборудования. Разработчики рассматриваемой измерительной системы взяли за основу оборудование компании ADVANTECH, что позволило им создать систему большой канальности с синхронизацией получаемой экспериментальной информации при минимальной стоимости одного измерительного канала и с высокой универсальностью в обслуживании различных типов датчиков. Разработанная мобильная система применяется при проведении дорожных и полигонных испытаний на плавность хода, управляемость и устойчивость, а также может быть использована для проверки степени нагруженности элементов конструкции разрабатываемого изделия.

В системе учтена возможность использования как аналоговых, так и цифровых датчиков, а также многоканальность измерений. Напряжение питания измерительной аппаратуры выполнено от бортовой сети мобильной машины.

1. Требования к системе и ее технические характеристики

Для проведения дорожных испытаний мобильная измерительная система должна обладать следующими свойствами:

- работоспособностью в условиях повышенных вибраций и при различных температурных изменениях;
- многоканальностью;
- универсальностью, т. е. независимостью от типов подключаемых датчиков (с условием, что датчики имеют выход по напряжению не более ± 10 В);
- возможностью:
 - подключения к бортовой сети питания автомобиля;
 - ввода параметров задания режима испытаний;
 - накопления больших массивов измерительной информации;
 - сохранения результатов испытаний и просмотра файлов предыдущих испытаний;
 - тарировки датчиков;
- иметь режим удаленного управления натурными испытаниями.

Исходя из этих требований была разработана и изготовлена измерительная система [5] с техническими характеристиками, приведенными в таблице.

Основные технические характеристики мобильной измерительной системы для проведения дорожных испытаний

Параметр	Значение
Напряжение питания	12–36 В
Аналоговые входы	64 канала
– разрядность АЦП	16 бит
– частота выборки, макс.	1 кГц
Диапазон входных напряжений (в зависимости от усилителя)	$\pm 10V$; $\pm 5V$; $\pm 2,5V$; $\pm 1,25V$; $\pm 0,625V$
Поканально программируемый коэффициент усиления	0,5, 1, 2, 4 и 8
Флеш-диск	6 Гбайт
Автокалибровка, программируемая схема запуска, память FIFO: 1024 байта для каждого канала, рабочий диапазон температур 0–60 °С	
Возможность подключения тензорезисторов и тензодатчиков по четырехпроводной схеме подключения	
Возможность подключения датчиков с токовым выходом, потенциальным и термопар	
Габаритные размеры, мм (ДхШхВ)	600x220x550
Вес, кг	≈ 45

2. Описание функционирования системы

Мобильная измерительная система (рис. 1) выполнена в алюминиевом корпусе, который закрывается ключом для ограничения к ней доступа. На передней панели находится дисплей и клавиатура, а также кнопка запуска, разъем USB и индикатор работы. Сверху корпуса располагаются пружинные клеммные колодки для подключения аналоговых и цифровых датчиков. По бокам корпуса расположены ручки для переноски и разъем для подключения к источнику питания.



Рис. 1. Внешний вид мобильной измерительной системы для проведения дорожных испытаний

Система построена на базе встраиваемого безвентиляторного промышленного компьютера с установленными в него PCI-платами аналогового и цифрового ввода (рис. 2).

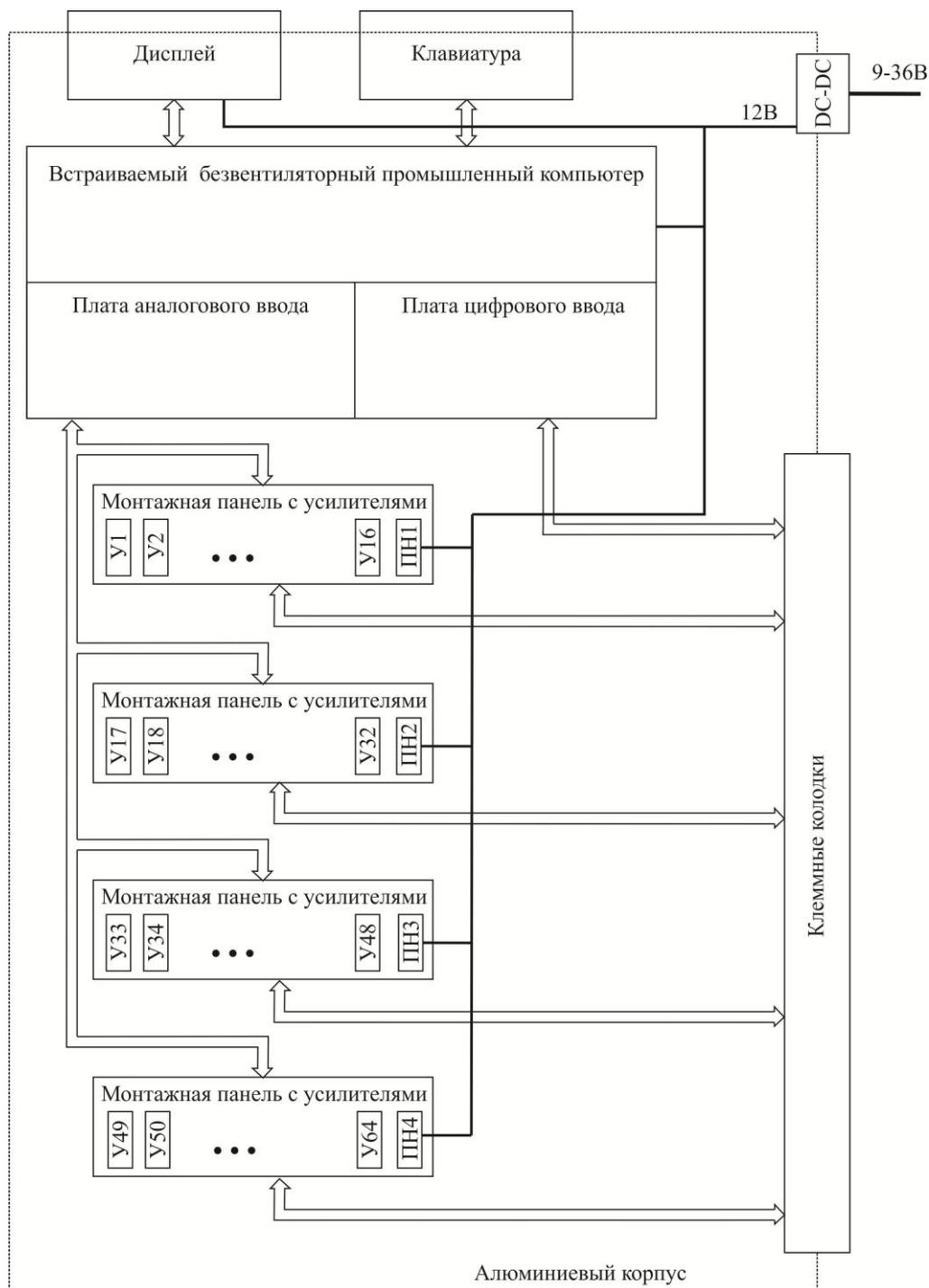


Рис. 2. Функциональная схема измерительной системы: У1-У64 – нормализатор сигнала; ПН1-ПН4 – преобразователь напряжения (12 В в +5 В); DC-DC – преобразователи напряжения (9–36 В в +12 В)

Промышленный компьютер содержит процессор Celeron 1,85 МГц, память 512 Мб, встроенный видеоконтроллер Intel 855 Graphics2 и различные интерфейсы: 2 PCI слота, 2×USB 2.0, 3×RS-232, 1×RS-232/422/485, 1×Parallel, 1×PS/2. Он изготовлен в алюминиевом корпусе с размерами 137x189x221 мм, весом 5 кг и температурным диапазоном от 0 до +60 °С. Промышленный компьютер имеет собственный преобразователь напряжения, позволяющий использовать напряжения питания в диапазоне от 9 до 32 В постоянного тока. В режиме дорожных испытаний среднеквадратическая вибрационная нагрузка

в полосе частот 5–500 Гц не должна превышать $5g \text{ м/с}^2$. Для защиты от скачков напряжения в бортовой сети автомобиля дополнительно установлен DC-DC-преобразователь, входное напряжение которого 12–36 В, а выходное – 12 В.

К промышленному компьютеру подключены платы аналогового и цифрового ввода. Плата аналогового ввода подсоединена к четырем шестнадцатиканальным стандартным монтажным панелям для установки нормализаторов (усилителей). Монтажные панели соединены с внешними клеммными колодками, находящимися на корпусе измерительной системы. Плата цифрового ввода непосредственно соединена с внешними клеммными колодками, расположенными на корпусе измерительной системы.

Измерительная система устанавливается в салоне кабины автотранспортного средства. Для снижения уровня вибрации она закрепляется на сиденье с использованием виброгасящего коврика. К системе подключаются все измерительные датчики (рис. 3) и подается напряжение питания 24 В от бортовой сети к специально выведенным контактам. Напряжение питания понижается с помощью DC-DC-преобразователя до +12 В, которые необходимы для работоспособности встраиваемого компьютера, монитора и нормализаторов. В монтажной панели нормализаторов установлен дополнительный преобразователь напряжения (ПН1-ПН4, см. рис. 2) из 12 В в +5 В для непосредственного их питания.

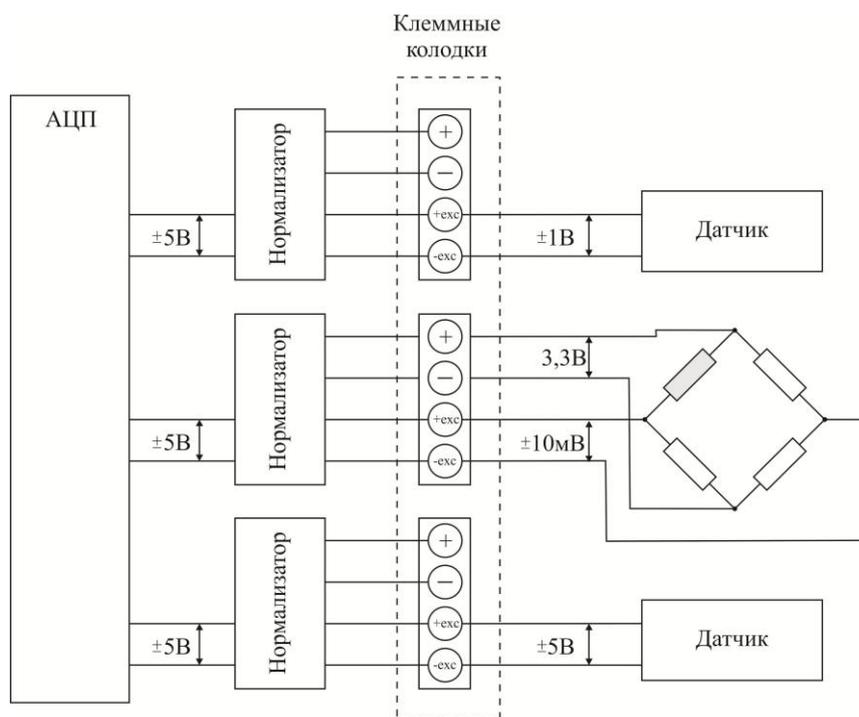


Рис. 3. Схемы подключения датчиков

3. Аппаратное обеспечение

В измерительной системе для подключения датчиков используются нормализаторы компании DATAFORTH, большое разнообразие которых позволяет подключать различные виды датчиков и гибко настраивать систему под проведение конкретных испытаний. В данной системе установлены нормализаторы с потенциальным входом $\pm 1\text{В}$ в количестве 32 шт., нормализаторы для тензорезисторов с входом $\pm 10\text{ мВ}$ в количестве 8 шт., остальные 24 канала замкнуты через монтажную панель, что позволяет сигналу от датчика непосредственно проходить без изменения в АЦП, т. е. выход датчика не должен превышать $\pm 5\text{ В}$. Всего в измерительной системе имеется 64 контактных разъема для подключения нормализаторов. Система

построена таким образом, что она не зависит от вида подключенных нормализаторов, дополнительной настройки (как аппаратной, так и программной) не требуется. В случае необходимости следует заменить нормализатор и провести тарировку датчика.

Все сигналы с нормализаторов подаются через клеммный адаптер и 68-контактный соединитель SCSI-II на плату аналогового ввода. Плата представляет собой 16-разрядное АЦП с максимальной частотой выборки до 250 кГц. Имеется 64 потенциальных или 32 дифференцированных канала. Кроме того, плата аналогового ввода имеет программируемый коэффициент усиления по каждому каналу и программируемую систему запуска измерений, что позволяет организовать синхронное измерение с датчиков.

Цифровые входы подключены через клеммный адаптер и 50-контактный соединитель SCSI-II к плате цифрового входа. Устройство представляет собой 64-канальную плату цифрового ввода с гальванической изоляцией, широким диапазоном входных сигналов (10–50 В постоянного тока) и возможностью работы по прерыванию. Платы поставляются с набором драйверов и библиотек, что значительно облегчает работу с ними.

Если подключить адаптеры 3G или Wi-Fi через разъем USB, появляется возможность работы с измерительной системой удаленно (рис. 4).

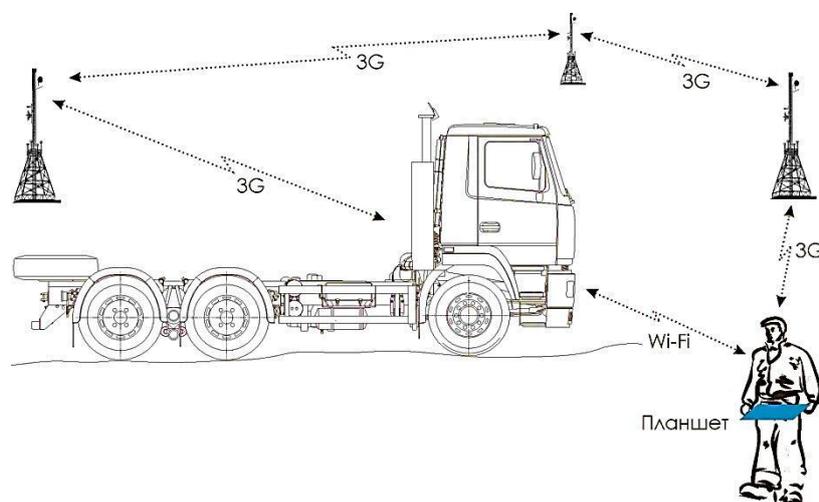


Рис. 4. Удаленное проведение испытаний

С помощью адаптера 3G (через Ethernet) оператор может, находясь в любой точке земного шара, проводить испытания (настраивать систему, получать, просматривать и сохранять результаты испытаний), используя мобильные устройства типа планшета. С помощью адаптера Wi-Fi оператор может управлять испытаниями с расстояния от 100 до 300 м в зависимости от места их проведения.

Единственным условием к мобильным устройствам для удаленных испытаний является наличие операционной системы Windows XP и выше, так как клиентская часть программы написана на Borland Delphi. Для работы с измерительной системой разработчики использовали планшет компании iRU с установленной на нем операционной системой Windows 7 и с адаптерами 3G и Wi-Fi.

4. Программное обеспечение

Программное обеспечение разработано в среде Borland Delphi и работает под управлением ОС Windows XP. Использование ОС Windows XP позволяет легко обучиться работе с системой и не требует специальной подготовки оператора. Данное программное обеспечение зарегистрировано в Национальном центре интеллектуальной собственности (свидетельство № 368 от 30.11.2011 г.).

Взаимодействие программы с аппаратной частью осуществляется с использованием библиотек, с помощью которых плата АЦП настраивается на режим синхронной работы.

Это позволяет получать информацию по всем каналам одновременно, что необходимо для исследования динамики объекта испытания.

Основная особенность программного обеспечения – это использование мультимедийного таймера для соблюдения точного шага по времени. Данный таймер позволяет выдерживать точный шаг квантования по времени с точностью до 1 мс при работе платы АЦП в синхронном режиме.

При проведении испытания программа работает параллельно по двум потокам: один управляет работой измерительных плат с сохранением измеренной информации, второй реализует пользовательский интерфейс (реагирует на команды пользователя).

Первый поток производит:

- настройку каналов измерительных плат (активности, коэффициента усиления);
- предварительную обработку измеренной информации (нормирование сигналов);
- накопление измеренной информации в памяти.

Второй поток осуществляет следующие операции:

- создание экранных форм;
- отображение измеренной информации;
- построение графиков с возможностью их просмотра по всем каналам или выборочно;
- спектральный анализ измеренной информации (построение амплитудно-частотной и фазочастотной характеристик);
- сохранение измеренной информации.

На рис. 5 и 6 показаны формы программы управления измерительной системы и окна просмотра результатов измерения и обработки.

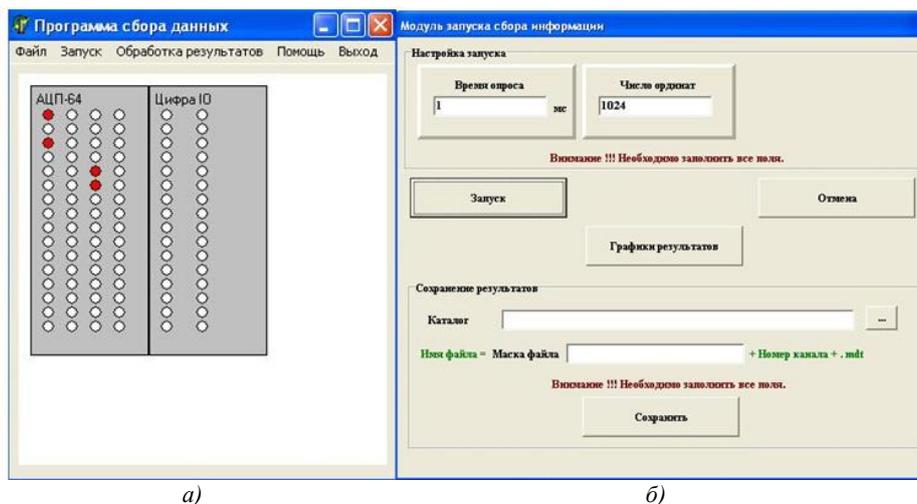


Рис. 5. Экранные формы программы измерительной системы: а) после настройки каналов измерения; б) после проведения измерений по каналам

Дружественный, простой и интуитивный интерфейс программы управления обеспечивает быструю настройку системы обработки под конкретные дорожные испытания.

Методы работы с использованием сети Wi-Fi и 3G принципиально не отличаются. Для управления измерительной системой используется любое бесплатное программное обеспечение для удаленного управления компьютером. Для работы в сети Wi-Fi использовалась программа TeamViewer, благодаря которой оператор получает доступ к компьютеру измерительной системы. После настройки соединения через программу TeamViewer оператор на экране планшета видит «клонированный экран» измерительной системы и имеет возможность удаленно осуществлять все действия, как будто находится непосредственно перед измерительной системой.

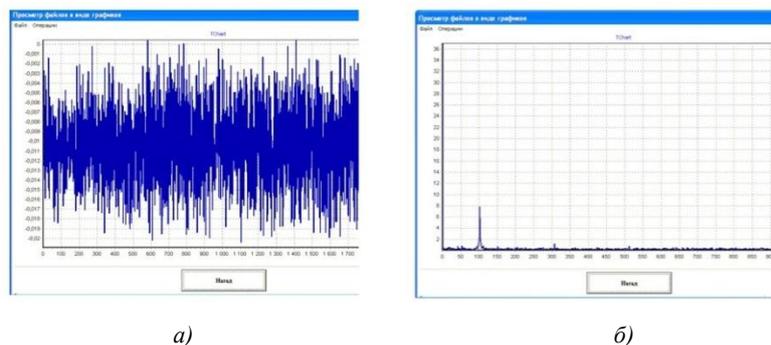


Рис. 6. Экранные формы программы измерительной системы: а) результатов измерения; б) обработки результатов измерения (амплитудно-частотная характеристика)

Заключение

Натурные испытания являются источником информации для создания верифицированных компьютерных моделей, что позволит оптимизировать изделие по различным его свойствам. Важную роль в проведении натурных испытаний играет измерительная система, которая является источником объективной информации о процессах, протекающих в исследуемом изделии.

Описанная мобильная измерительная система разработана и изготовлена в Республиканском компьютерном центре машиностроительного профиля Объединенного института машиностроения НАН Беларуси. Данная система используется для проведения натурных испытаний многозвенных автопоездов на плавность хода, устойчивость и маневренность.

Разработанная измерительная система обладает рядом свойств, позволяющих проводить натурные испытания на высоком уровне:

- многоканальностью, обеспечивающей одновременное измерение от различных видов датчиков, что значительно сокращает время проведения натурных испытаний;
- универсальностью, обуславливающей большой выбор подключаемых датчиков в диапазоне ± 10 В, что также сокращает затраты времени на перекомпоновку и переналадку измерительной системы;
- синхронностью, благодаря которой достигается высокая точность анализа динамических характеристик исследуемого объекта;
- удаленным управлением, позволяющим не только управлять процессом натурных испытаний на расстоянии, но и проводить наладку измерительной системы, обработку и передачу экспериментальной информации.

Все эти свойства измерительной системы свидетельствуют о ее большой практической ценности.

Список литературы

1. Технология моделирования динамики колебаний грузовых автомобилей с использованием стендовых и виртуальных испытаний / В.С. Кончак [и др.] // Весці НАН Беларусі. Сер. фіз.-тэхн. навук. – 2011. – № 3. – С. 83–91.
2. Цифровая автоматизированная система тестирования объектов испытаний, анализа и обработки экспериментальных данных / П.М.Чеголин [и др.] // Автоматизация конструкторско-технологического проектирования : сб. науч. тр. – Минск : Ин-т техн. кибернетики АН БССР, 1985. – С. 4–11.
3. Система КАМАК. Последовательная магистраль интерфейсной системы : ГОСТ 26.201.2-94. – Введ. 01.07.96. – Минск : Межгос. совет по стандартизации, метрологии и сертификации : Белорус. гос. ин-т стандартизации и сертификации, 1995. – 72 с.
4. Система КАМАК. Организация многокрейтовых систем. Требования к магистрали ветвей и крейт-контроллеру КАМАК типа А1 : ГОСТ 26.201.1-94. – Введ. 01.07.96. – Минск : Межгос. совет по стандартизации, метрологии и сертификации : Белорус. гос. ин-т стандартизации и сертификации, 1996. – 28 с.

5. Мобильная многоканальная система для проведения стендовых и натурных испытаний / В.С. Кончак [и др.]. // Тез. докл. 7-й Междунар. конф. «Информационные технологии в промышленности». – Минск : ОИПИ НАН Беларуси, 2012. – С. 151–152.

Поступила 18.11.2013

*Объединенный институт
машиностроения НАН Беларуси,
Минск, ул. Академическая, 12
e-mail: krom_hsv@bk.ru*

S.V. Hitrikov

**MEASURING MOBILE MULTI-CHANNEL SYSTEM
FOR BENCH AND TRACK TESTS**

A measuring system is designed for full-scale testing of units and aggregates of mobile machines in order to verify their compliance with the standards defined by the product specification. The system is designed and manufactured using modern components, which provide unique properties: multi-channel, versatility, synchronicity and remote management.

УДК 004.725

Ю.И. Воротницкий, В.П. Кочин, Д.А. Стрикелев

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ ДЛЯ ОПТИМИЗАЦИИ СТРУКТУРЫ БЕСПРОВОДНОЙ СЕТИ С ЗАДАНЫМ КАЧЕСТВОМ ОБСЛУЖИВАНИЯ

Описывается математическая модель оптимизации расположения точек доступа Wi-Fi с заданным качеством обслуживания. Разрабатывается модифицированный генетический алгоритм оптимизации расположения точек доступа Wi-Fi. Проводится апробация результатов для тестового здания.

Введение

Задачи проектирования и оптимизации сетей передачи данных, актуальные на данный момент в науке и промышленности, характеризуются высокой вычислительной сложностью и немонотонным, «зашумленным» ландшафтом пространства решений. При решении этих задач традиционные аналитические методы оптимизации либо неприменимы, либо находят локальные субоптимальные решения, далекие от глобальных экстремумов, а методы полного перебора неприменимы из-за размерности пространства решений. Это обусловило появление и развитие нового класса эвристических методов, основанных на имитации механизмов, действующих в природе, в частности генетических алгоритмов. Назовем их наиболее важные преимущества:

- высокая степень распараллеливаемости на различных уровнях (как микро-, так и макро-);
- эффективность при решении многокритериальных задач по следующим причинам: поиск ведется в различных областях пространства решений одновременно, не требуется априорного задания весов критериев, на выходе получаются множества решений с различными приоритетами критериев.

1. Постановка задачи

Требуется определить оптимальное расположение точек доступа Wi-Fi в заданном помещении с целью организации беспроводного доступа к мультимедийным ресурсам с требуемым качеством обслуживания и с учетом особенностей организации учебного процесса вуза – требований на возможность подключения заданного количества пользователей в аудиториях [1].

Размещение пользователей в помещении задается выражением

$$U = \{x, y\}_i, \quad i = \overline{1, R}, \quad (1)$$

где R – общее число пользователей; x, y – координаты пользователей.

Подключение пользователей к точкам доступа задано матрицей UM размерностью $R \times N$:

$$UM_{ij} = \begin{cases} 1, & \text{если } i\text{-й пользователь подключен к } j\text{-й точке доступа;} \\ 0, & \text{если } i\text{-й пользователь не подключен к } j\text{-й точке доступа,} \end{cases} \quad (2)$$

где R – количество пользователей; N – количество точек доступа.

Помещение представляет собой множество стен с заданными координатами их углов, толщиной и проницаемостью материала стены

$$W_i = \{f_i(x); x'_i; x''_i; h_i; \varepsilon_i\}, \quad i = \overline{1, M}, \quad (3)$$

где M – количество стен в здании; $f_i(x)$ – уравнение i -й стены с граничными условиями x'_i, x''_i ; h_i – толщина i -й стены; ε_i – диэлектрическая проницаемость стены.

Возможность подключения пользователей к точке доступа определяется ее техническими характеристиками, а также мощностью сигнала, ослабляемого при прохождении препятствий.

Для расчета ослабления сигнала применяется выражение, полученное в [2] и зависящее от следующих аргументов:

d – расстояния в свободном пространстве, которое проходит волна;

h – толщины стены, через которую проходит волна;

k – коэффициента ослабления материала стены;

f – частоты волны.

С помощью выражения [2] становится возможным вычислить ослабление сигнала при распространении из точки (x_i, y_i) в точку (x'_i, y'_i) :

$$\Psi(x_i; y_i; x'_i; y'_i). \quad (4)$$

Для расчета скорости доступа при заданной мощности сигнала и количестве пользователей применяется выражение, приведенное в [3]. На основе данного выражения и требований качества сервиса положим:

$$\varphi(P_i) = \begin{cases} 1, & \text{если } i\text{-й пользователь подключен с соблюдением требований } QoS; \\ 0 & \text{в противном случае.} \end{cases} \quad (5)$$

Таким образом, целевая функция оптимизации приобретет вид

$$\begin{cases} \sum_{i=1}^R \sum_{j=1}^N UM_{ij} \times \phi(\Psi(x_i; y_i; x'_j; y'_j)) \rightarrow \max; \\ \sum_{i=1}^N UM_{ij} \leq 1, \forall i \in \overline{1, R}, \end{cases} \quad (6)$$

где x_i, y_i – координаты i -го пользователя; x'_j, y'_j – координаты j -й точки доступа Wi-Fi.

Искомым решением задачи является множество координат точек доступа X (размерности N), $X_i = \{x'_i, x''_i\}$.

2. Генетический алгоритм оптимизации размещения точек доступа

В рамках описанной математической модели задачи решением будет являться набор координат точек доступа. Координаты каждой точки будут определяться геном, а их множество – формировать хромосому.

Первым шагом алгоритма является генерация начальной популяции решений. В классических генетических алгоритмах начальная популяция генерируется случайным образом. Однако в рассматриваемом случае это может привести к росту вычислительной сложности и к вырождению конечного решения. В связи с этим предлагается двухэтапный алгоритм формирования исходной популяции. На первом этапе необходимо определить места с наибольшим количеством пользователей Wi-Fi и центры этих «скоплений». Для решения подобных задач исследователями применялись такие алгоритмы кластеризации, как K-means, C-means, Fuzzy-clustering и др. [4–6].

Решаемая задача характеризуется рядом ограничений:

- количеством пользователей, подключаемых к точке доступа;
- геометрией помещения (возможно, ближайший клиент будет находиться через стенку, но по мощности будет самый худший);
- вычислительной сложностью, с учетом того что необходимо сгенерировать популяцию с Q количеством особей.

Таким образом, ни один из перечисленных выше алгоритмов не удовлетворяет условиям задачи. В связи с этим в данной статье предложен следующий алгоритм (рис. 1).

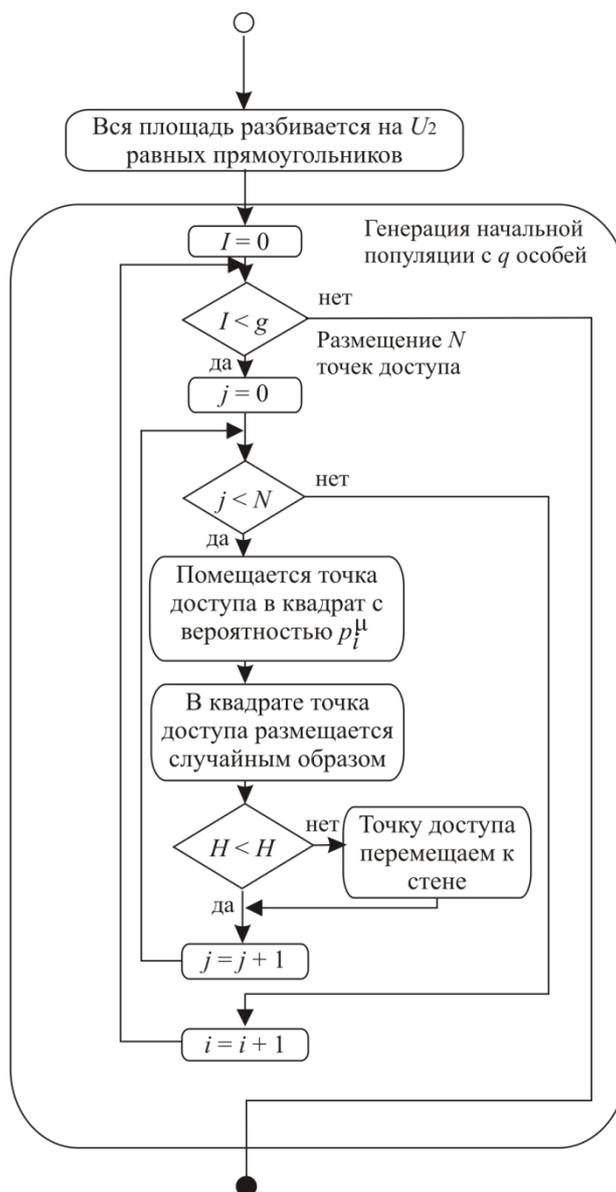


Рис. 1. Схема алгоритма генерации начальной популяции

Геометрическая конфигурация помещения, в котором размещаются точки доступа, вписывается в прямоугольник с координатами левого нижнего угла $(x_{min}; y_{min})$ и правого верхнего угла $(x_{max}; y_{max})$.

Вся площадь расположения точек доступа Wi-Fi разбивается на U^2 квадратов. При этом U определяется соотношением

$$(U - 1)^2 < N \leq U^2, \tag{7}$$

где N – количество точек доступа Wi-Fi, которые необходимо разместить.

Далее определяется количество абонентов в каждом квадрате, для чего точки доступа помещаются в квадрат с вероятностью p_i^u , пропорциональной количеству абонентов:

$$p_i^u = \frac{R_i}{R}, \tag{8}$$

где R_i – количество абонентов в i -м квадрате; R – общее количество абонентов.

На втором этапе работы алгоритма точка доступа случайным образом помещается внутри i -го квадрата. Также необходимо учитывать, что осуществить установку точки доступа возле

стены технически более просто. Для учета данного допущения определяется расстояние H до ближайшей стены, и если $H < H_0$, где H_0 – эталонное расстояние равное 0,5 м, то точку доступа перемещаем к стене по перпендикуляру.

На следующем этапе определяется алгоритм подключения клиентов к точке доступа (рис. 2):

1. Если $R^k < R_{max}^k$, где R^k – количество подключенных клиентов к точке доступа k , R_{max}^k – максимально разрешенное количество абонентов точки доступа k , то переходим к шагу 2; иначе, если не существует точки доступа с $R^k < R_{max}^k$, устанавливаем скорость доступа для клиента в точке $X B = 0$.

2. Для клиента в точке X рассчитываем скорость доступа B_k при подключении к точке доступа K .

3. Подключаем клиента к точке доступа с $\max(B_k)$.

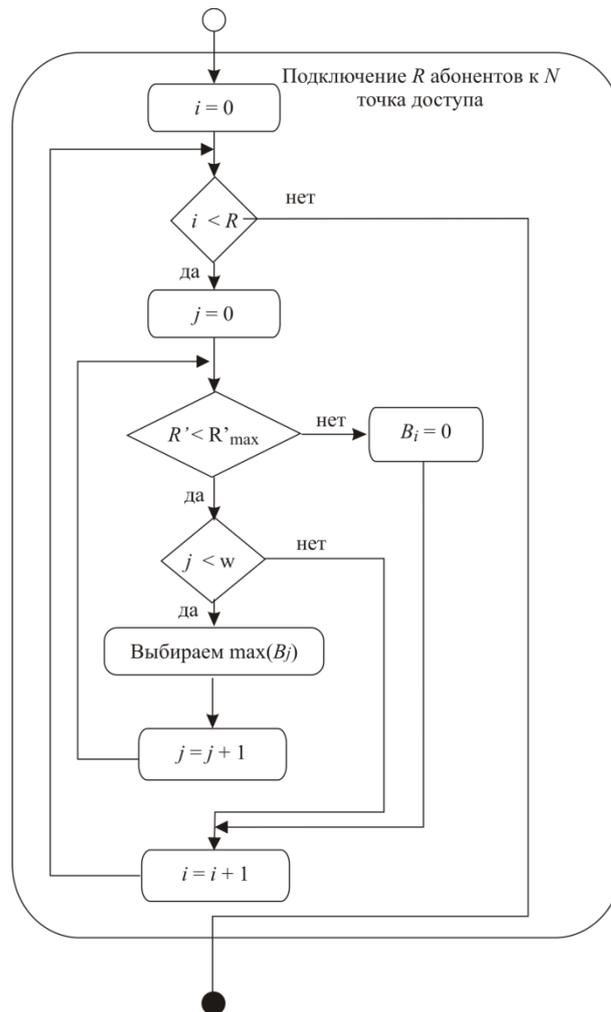


Рис. 2. Схема алгоритма подключения клиентов к точке доступа

Оценивание приспособленности хромосом в популяции состоит в расчете функции приспособленности (фитнес-функции) для каждой хромосомы этой популяции:

$$F = \sum_i^{b_i < B} c_i, \quad (9)$$

где B – пороговая скорость подключения клиентов к сети, определяемая моделью трансляции качества услуг QoS; b_i – скорость подключения i -го клиента к сети; c_i – штрафная функция для точки доступа, для которой $b_i < B$. При этом в соответствии с поставленной задачей вид штраф-

ной функции для точки доступа k , для которой выполняется условие $b_i < B$, должен зависеть от b_i и удовлетворять соотношению

$$\sum_{i \neq k} c_i^{\max} < \sum_i c_i^{\min}. \quad (10)$$

При рассмотрении качества услуг (Quality of Service, или QoS) в зависимости от конкретного типа решаемой задачи начальные метрики критерия формулируются с точки зрения пользователя на пользовательском уровне. В качестве характеристик прикладного уровня используются время задержки при передаче файла и характеристики потока. В качестве характеристик транспортно-сетевого уровня используются пропускная способность, время задержки, джиттер, вероятность потери пакета. В качестве характеристики физического уровня используется мощность сигнала. Так как характеристика физического уровня является непосредственно измеряемой (оцениваемой) в процессе оптимизации, авторы предлагают последовательное использование трех преобразований, каждое из которых выполняется по своим правилам. На первом этапе необходимо осуществить преобразование характеристик пользовательского уровня в характеристики прикладного уровня, на втором этапе – преобразование характеристик прикладного уровня в характеристики транспортно-сетевого уровня, на третьем этапе – преобразование характеристик транспортно-сетевого уровня в характеристики физического уровня.

Таким образом, штрафная функция для точек доступа будет иметь вид

$$c_i = \text{const} + e^{B-b_i}, \quad (11)$$

где const – константа, которая зависит от количества точек доступа, участвующих в оптимизации, и параметра B , который определяется моделью трансляции качества услуг QoS:

$$\text{const} \geq R(e^B - 1), \quad (12)$$

Фитнес-функция будет иметь вид

$$F = \sum_i^{b_i < B} (\text{const} + e^{B-b_i}). \quad (13)$$

В качестве оператора мутации выступает одно из следующих преобразований:

- удаление точки доступа с вероятностью α ;
- включение точки доступа с вероятностью β . Смотрим, какие точки доступа выключены, случайно выбираем одну, включаем ее в точке со случайной координатой (генерация случайных координат должна попадать в разрешенную координату);
- смещение с вероятностью γ точки доступа Wi-Fi в новое положение.

На первом этапе для скрещивания случайным образом выбираются пары хромосом из родительской популяции. Это временная популяция, состоящая из хромосом, отобранных в результате селекции и предназначенных для дальнейших преобразований операторами скрещивания и мутации с целью формирования новой популяции потомков. Само скрещивание происходит по следующим правилам:

1. Для пары хромосом вычисляются вероятности $p_1^c = \frac{F_1}{F_1 + F_2}$, $p_2^c = 1 - p_1^c$, где F_1 и F_2 – значения фитнес-функции для двух хромосом.

2. Далее проверяются следующие утверждения:

- точка доступа присутствует в первой хромосоме, но отсутствует во второй. Тогда с вероятностью p_1^c она присутствует в потомке, с вероятностью p_2^c не присутствует;
- точка доступа присутствует во второй хромосоме, но отсутствует в первой. Тогда с вероятностью p_2^c она присутствует в потомке, с вероятностью p_1^c не присутствует;
- точка доступа присутствует в обеих хромосомах. Тогда координаты точки доступа потомка берутся с вероятностью p_1^c из первой хромосомы, с вероятностью p_2^c из второй хромосомы.

После скрещивания размер популяции является увеличенным по отношению к своему нормальному уровню. Для корректировки этой ситуации на финальном этапе итерации проводится отбор наиболее приспособленных индивидов на основе значений функции приспособленности. Такой выбор производится согласно принципу естественного отбора, по которому наибольшие шансы на участие в создании новых особей имеют хромосомы с наименьшими значениями функции приспособленности. Таким образом, в алгоритме будет использована турнирная селекция, при которой в результате скрещивания пары хромосом будет выбрана особь с наилучшим значением функции приспособленности.

На основе модифицированного генетического алгоритма был создан программный комплекс оптимизации расположения точек доступа Wi-Fi. Для создания программного комплекса использовались язык программирования C++ и библиотека MPI.

Для проведения вычислительного эксперимента предполагалось, что необходимо разместить четыре точки доступа Wi-Fi в помещении с девятью комнатами, одна из которых поточная аудитория, при этом количество пользователей в комнатах известно.

Расчет расположения точек доступа Wi-Fi производился на компьютере со следующими параметрами: операционная система Windows 7, двухъядерный процессор Intel Pentium Dual E2160 1,8 ГГц, оперативная память 2048 Мб.

Для отображения результатов вычисления был создан графический интерфейс. В процессе вычисления проводилась выборка хромосом для оценки сходимости алгоритма. Результат вычисления показан на рис. 3.

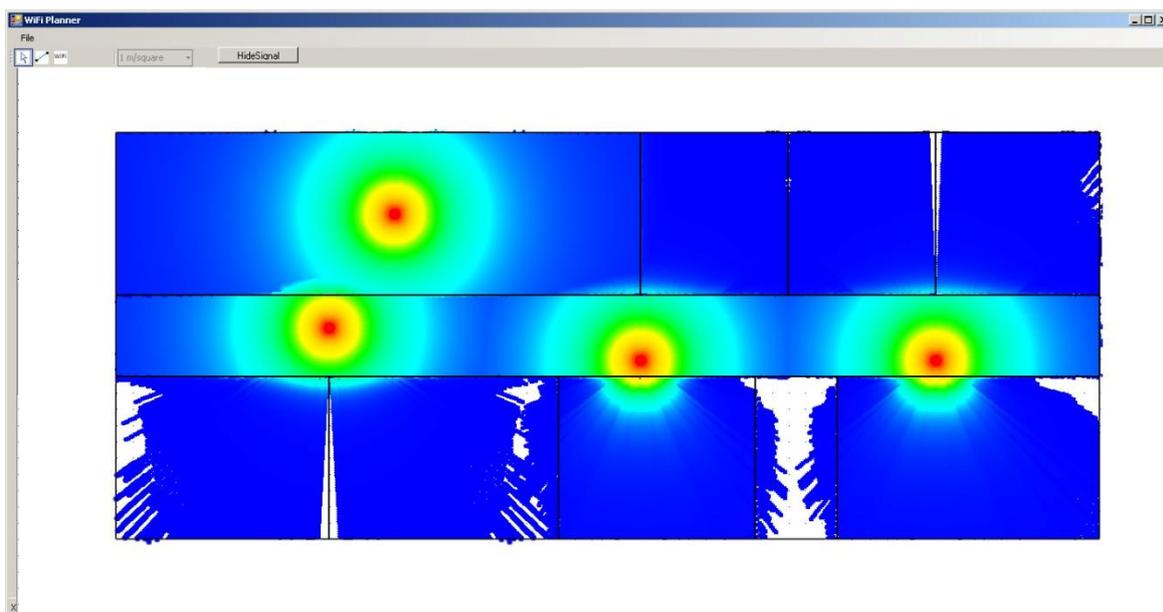


Рис. 3. Размещение точек доступа Wi-Fi с наилучшим значением фитнес-функции

Среднее время работы программного комплекса составило 915 с при размере популяции 60 особей.

В результате проведенного эксперимента удалось установить следующие закономерности:

- при увеличении количества точек доступа W-Fi размер популяции увеличивается по полиномиальному закону;
- при увеличении количества точек доступа W-Fi время работы алгоритма увеличивается по экспоненциальному закону.

Данные зависимости объясняются многократным вычислением целевой функции на каждой итерации. Таким образом, при увеличении количества точек доступа Wi-Fi до 50 время работы алгоритма может составить до года в зависимости от размера популяции и геометрии по-

мещения. В связи с этим в дальнейшем вычисления будут проводиться с использованием вычислительных ресурсов суперкомпьютера СКИФ-БГУ. В настоящее время для расчетов доступны 65 узлов, в каждом из которых два двухъядерных процессора AMD Opteron 2,2 ГГц. Таким образом, скорость работы алгоритма увеличится примерно в 300 раз по сравнению с вычислениями на тестовой рабочей станции.

Использование ресурсов суперкомпьютера СКИФ-БГУ позволит, во-первых, рассчитать оптимальное расположение точек доступа Wi-Fi при их достаточно большом количестве, во-вторых, увеличить точность полученных результатов.

Заключение

Предложенный генетический алгоритм позволяет получить размещение точек доступа Wi-Fi при условии наиболее полного соблюдения требований заданного качества услуг. Данный алгоритм характеризуется высокой степенью распараллеливаемости, что позволяет при расчетах использовать ресурсы суперкомпьютерных кластерных систем.

Список литературы

1. Воротницкий, Ю.И. Мобильные компьютерные устройства в «облачной» информационно-образовательной среде общеобразовательной школы / Ю.И. Воротницкий, М.Г. Зеков, А.Н. Курбацкий. – Минск : Ривш, 2012. – 100 с.
2. Кочин, В.П. Быстрая оценка мощности Wi-Fi-сигнала при прохождении препятствий в пределах здания / В.П. Кочин, Ю.И. Воротницкий, Д.А. Стрикелев // Вестник БГУ. Сер. 1. – 2013. – № 1. – С. 45–50.
3. Воротницкий, Ю.И. Оценка пропускной способности информационного канала беспроводной связи с заданной мощностью сигнала / Ю.И. Воротницкий, В.П. Кочин, Д.А. Стрикелев // Информатизация образования. – 2013. – № 3. – С. 28–32.
4. Jain, A. Data Clustering: A Review / A. Jain, M. Murty, P. Flynn // ACM Computing Surveys. – 1999. – Vol. 31, no. 3. – P. 264–323.
5. Прикладная статистика: классификация и снижение размерности / С.А. Айвазян [и др.]. – М. : Финансы и статистика, 1989. – 607 с.
6. Gorban, A.N. Principal Graphs and Manifolds. Ch. 2 / A.N. Gorban, A.Y. Zinovyev // Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques. – Hershey, PA, USA, 2009. – P. 28–59.

Поступила 12.03.2014

*Белорусский государственный университет,
Минск, пр. Независимости, 4
e-mail: kochyn@bsu.by*

Y.I. Vorotnitsky, V.P. Kochyn, D.A. Strikelev

A GENETIC ALGORITHM FOR WI-FI NETWORK STRUCTURE OPTIMIZATION UNDER THE GIVEN QUALITY OF SERVICE

A mathematical model for optimal placement of Wi-Fi access points under the given quality of service is developed. A modified genetic algorithm for optimal placement of Wi-Fi access points is worked out. Probation of the results is carried out for a test area.

ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ

УДК 004.33.054

В.Н. Ярмолик, С.В. Ярмолик

АДРЕСНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ
ДЛЯ МНОГОКРАТНОГО ТЕСТИРОВАНИЯ ОЗУ

Предлагается универсальный метод генерирования адресных последовательностей с заданными свойствами для многократных маршевых тестов оперативных запоминающих устройств. В качестве математической модели используется модификация экономичного способа Антонова и Салеева для формирования последовательностей Соболя. В рамках предлагаемой модели последовательности Соболя являются подмножеством адресных последовательностей, наряду с которыми формируются последовательности кода Грея, анти-Грея, пересчетные и ряд других последовательностей, в том числе последовательности с заданными свойствами.

Введение

В настоящее время актуальной является проблема тестирования современных вычислительных систем, таких как встроенные системы (Embedded Systems), системы на кристалле (Systems-on-a-Chip) и сети на кристалле (Nets-on-a-Chip) [1, 2]. Характерной особенностью подобных вычислительных систем является неуклонное увеличение емкости их запоминающих устройств, и в первую очередь оперативных запоминающих устройств (ОЗУ) [1–3]. Соответственно возрастает актуальность их эффективного тестового диагностирования [4, 5]. Тестирование современных ОЗУ с целью обнаружения различных неисправностей предполагает запись в ОЗУ всевозможных его состояний и их считывание, что приводит к нереально большой сложности теста, так как сложность самой тестовой процедуры пропорциональна величине 2^N , где N – емкость ОЗУ в битах. Поэтому в настоящее время используются и по-прежнему разрабатываются методы построения тестов, которые имеют существенно меньшую сложность, как правило, линейно зависящую от емкости ОЗУ [6, 7]. Эти тесты имеют общее название «маршевые тесты» (March Tests) и доминируют в практических приложениях [8]. При их применении необходимо использование небольшого количества тестовых процедур, которые тем не менее были бы достаточно качественными, т. е. покрывали максимально возможное число различных неисправных состояний ОЗУ, и в первую очередь сложных неисправных состояний [6–8].

Первоначально методы обнаружения сложных неисправностей ОЗУ рассматривались для случая разрушающих тестов [9], т. е. таких тестов, применение которых приводит к потере информации, хранимой в ОЗУ. Следующим шагом в развитии методов тестирования ОЗУ явилось использование неразрушающих тестов [10]. При многократном применении неразрушающего теста для различных состояний запоминающих ячеек ОЗУ в силу перераспределения подмножеств обнаруживаемых и необнаруживаемых неисправностей оказывается возможным достижение 100 %-й полноты покрытия неисправностей ОЗУ [10, 11]. Многократное применение маршевых тестов с различными адресными последовательностями явилось следующим этапом развития неразрушающих тестов, что позволило существенно повысить эффективность тестирования ОЗУ [12, 13]. Однако вопросы выбора оптимальных адресных последовательностей для многократного тестирования ОЗУ, их соотношение и численные оценки эффективности тестирования ОЗУ находятся лишь на начальной стадии исследований [4, 12]. Одной из основных проблем многократного тестирования ОЗУ является ограниченность набора возможных адресных последовательностей и сложность алгоритмов их генерирования [4, 7, 11–13].

В предлагаемой статье рассматривается универсальный метод генерирования широкого набора адресных последовательностей для реализации многократных маршевых тестов ОЗУ.

1. Адресные последовательности

Под адресной последовательностью понимают упорядоченную последовательность m -разрядных двоичных векторов $A(n) = a_{m-1}a_{m-2}a_{m-3}\dots a_2a_1a_0$, где $a_i \in \{0, 1\}$, $i \in \{0, 1, 2, \dots, m-1\}$, и $n \in \{0, 1, 2, \dots, 2^m-1\}$, однократно принимающих все возможные значения из множества возможных адресов $\{0, 1, 2, \dots, 2^m-1\}$. Адресная последовательность удовлетворяет следующему основному свойству [4, 7, 12].

Свойство 1. Адресная последовательность $A(n) = a_{m-1}a_{m-2}a_{m-3}\dots a_2a_1a_0$ состоит из 2^m возможных m -разрядных двоичных комбинаций, формируемых в произвольном порядке.

Подобные последовательности имеют период, равный 2^m , и называются последовательностями максимальной длины, или последовательностями де Брюйна (de Bruijn sequences) [4, 7, 11, 12]. Общее количество адресных последовательностей, равное

$$2^m! \approx (2^m)^{2^m} e^{-2^m} \sqrt{2\pi 2^m} \quad (1)$$

для заданного значения m и адресного пространства, состоящего из 2^m адресов, принимает астрономические значения.

Среди всевозможных адресных последовательностей выделяют три основных множества. К первому множеству относятся различные *детерминированные* последовательности, включая пересчетные (счетчиковые) последовательности, последовательности Грея, анти-Грея, последовательности с максимальным расстоянием Хэмминга и ряд других [7, 11–13]. Ко второму множеству относятся *псевдослучайные* последовательности, которыми называются последовательности неслучайных чисел, обладающие свойствами случайных последовательностей [2]. К третьему множеству относятся так называемые *квазислучайные* последовательности [2, 4, 14–17]. Такие последовательности в русскоязычной литературе называют ЛПт-последовательностями [15, 17]. Это означает, что любой последовательный участок последовательности хорошо распределен (более равномерно по сравнению с псевдослучайными последовательностями). В англоязычной литературе их называют последовательностями с малым дискрепансом (Low-Discrepancy Sequence), а их разновидности – по именам авторов [2, 14–16].

Наибольший интерес для целей многократного тестирования вызывают детерминированные последовательности с различными свойствами, которые оцениваются численными характеристиками, такими как расстояние Хэмминга, расстояние Минковского и его модификации [7, 18]. Как показано в [18], выбор адресных последовательностей для обеспечения высокой эффективности тестирования ОЗУ целесообразно осуществлять на основе расстояния Минковского, которое для $\lambda = 1$ принимает вид арифметического расстояния (Манхэттенского расстояния).

2. Математическая модель генерирования адресных последовательностей

Последовательности Соболя (Sobol) используют двоичную систему счисления для формирования координат точек в s -мерном пространстве и являются широко востребованными для современных приложений [14–17, 19, 20], в том числе для генерирования адресных последовательностей $A(n)$ [2, 4].

В общем случае значение координат n -го элемента последовательности Соболя вычисляется как поразрядная сумма по модулю два до $m = \lfloor \log_2 n \rfloor$ операндов в зависимости от количества ненулевых компонент двоичного представления $b_{m-1}(n)b_{m-2}(n)\dots b_1(n)b_0(n)$ числа n . Количество операндов может быть снижено до одного при использовании экономичного способа Антонова и Салеева, основанного на представлении числа n в коде Грея [2, 4, 21]. Тогда формирование n -го элемента $A(n)$ последовательности Соболя осуществляется в соответствии с соотношением

$$A(n) = A(n-1) \oplus v_i, \quad n = \overline{0, 2^m - 1}, \quad i = \overline{0, m-1}, \quad (2)$$

в котором к предыдущему элементу $A(n-1)$ последовательности Соболя добавляется только одно модифицированное направляющее число $v_i, i \in \{0, 1, 2, \dots, m-1\}$ [2]. Значение индекса i направляющего числа v_i в выражении (2) зависит от так называемой последовательности переключений T_{m-1} отраженного кода Грея [22]. Для $m = 4$ эта последовательность имеет вид $T_3 = 0, 1, 0, 2, 0, 1, 0, 3, 0, 1, 0, 2, 0, 1, 0$. Формально последовательность переключений T_{m-1} определяет индекс i изменяемого разряда при переходе от n_g к n_{g+1} , где индекс g числа n_g означает представление в коде Грея исходного числа $n = b_{m-1}(n)b_{m-2}(n)\dots b_1(n)b_0(n)$. Число n в коде Грея может быть получено согласно известному соотношению $n_g = g_{m-1}(n)g_{m-2}(n)\dots g_1(n)g_0(n) = b_{m-1}(n)b_{m-2}(n)\dots b_1(n)b_0(n) \oplus 0b_{m-1}(n)b_{m-2}(n)\dots b_2(n)b_1(n)$ [22]. Сумма по модулю два последовательных значений n_g и n_{g+1} кода Грея и определяет индекс i направляющего числа v_i , используемого в выражении (2), последовательность значений которого и представляет собой T_{m-1} . Процедура получения последовательности переключений T_{m-1} для $m = 4$ приведена в табл. 1.

Таблица 1

Процедура генерирования последовательности переключений T_3

n	$n = b_3(n)b_2(n)b_1(n)b_0(n)$	$n_g = g_3(n)g_2(n)g_1(n)g_0(n)$	$n_g \oplus n_{g-1}$				T_3
			3	2	1	0	
0	0 0 0 0	0 0 0 0	0	0	0	0	
1	0 0 0 1	0 0 0 1	0	0	0	1	0
2	0 0 1 0	0 0 1 1	0	0	1	0	1
3	0 0 1 1	0 0 1 0	0	0	0	1	0
4	0 1 0 0	0 1 1 0	0	1	0	0	2
5	0 1 0 1	0 1 1 1	0	0	0	1	0
6	0 1 1 0	0 1 0 1	0	0	1	0	1
7	0 1 1 1	0 1 0 0	0	0	0	1	0
8	1 0 0 0	1 1 0 0	1	0	0	0	3
9	1 0 0 1	1 1 0 1	0	0	0	1	0
10	1 0 1 0	1 1 1 1	0	0	1	0	1
11	1 0 1 1	1 1 1 0	0	0	0	1	0
12	1 1 0 0	1 0 1 0	0	1	0	0	2
13	1 1 0 1	1 0 1 1	0	0	0	1	0
14	1 1 1 0	1 0 0 1	0	0	1	0	1
15	1 1 1 1	1 0 0 0	0	0	0	1	0

Значения разрядов кода Грея для $m = 4$ определяются в соответствии с соотношениями $g_3(n) = b_3(n), g_2(n) = b_2(n) \oplus b_3(n), g_1(n) = b_1(n) \oplus b_2(n), g_0(n) = b_0(n) \oplus b_1(n)$, полученными согласно выражению $n_g = g_3(n)g_2(n)g_1(n)g_0(n) = b_3(n)b_2(n)b_1(n)b_0(n) \oplus 0b_3(n)b_2(n)b_1(n)$ [22].

Конкретный вид модифицированной последовательности Соболя зависит от так называемых направляющих чисел (direction numbers) $v_i = \beta_{m-1}(i)\beta_{m-2}(i)\dots\beta_0(i), i \in \{0, 1, 2, \dots, m-1\}$ [2].

Необходимо отметить, что модифицированные направляющие числа v_i в силу ограничений на направляющие числа классических последовательностей Соболя также принимают определенные ограничения [2, 4, 17]. Так, для всех возможных направляющих чисел некоторые их разряды принимают фиксированные значения. Всегда $\beta_{m-1-i}(i) = 1, i = \overline{0, m-1}$, и $\beta_{m-1-j}(i) = 0$ для $j > i$, а $\beta_{m-1-i}(i)$ для $j < i$ принимают произвольные значения. Это значит, что для всех возможных последовательностей Соболя $v_0 = 100\dots 00, v_1 = \beta_{m-1}(1)10\dots 00$, где $\beta_{m-1}(1)$ принимает значение 0 либо 1, $v_2 = \beta_{m-1}(2)\beta_{m-2}(2)10\dots 00$ и т. д. [2]. В общем случае числа v_i можно представить в виде нижней треугольной матрицы с единичной диагональю (табл. 2) [4].

Таблица 2

Значения модифицированных направляющих чисел v_i

v_i	$\beta_{m-1}(i)$	$\beta_{m-2}(i)$	$\beta_{m-3}(i)$...	$\beta_1(i)$	$\beta_0(i)$
v_0	1	0	0	...	0	0
v_1	$\beta_{m-1}(1)$	1	0	...	0	0
v_2	$\beta_{m-1}(2)$	$\beta_{m-2}(2)$	1	...	0	0
...
v_{m-2}	$\beta_{m-1}(m-2)$	$\beta_{m-2}(m-2)$	$\beta_{m-3}(m-2)$...	1	0
v_{m-1}	$\beta_{m-1}(m-1)$	$\beta_{m-2}(m-1)$	$\beta_{m-3}(m-1)$...	$\beta_1(m-1)$	1

Соотношение (2) совместно с табл. 2 является обобщенной математической моделью одномерной последовательности Соболя (адресной последовательности), задаваемой m направляющими числами $v_i = \beta_{m-1}(i)\beta_{m-2}(i)\dots\beta_0(i)$, $i \in \{0, 1, 2, \dots, m-1\}$. Действительно, как показано в [2], набор значений $\beta_{m-1-i}(i) \in \{0, 1\}$ для $j < i$ определяет одну из $q = 2^{(m^2-m)/2}$ последовательностей Соболя. Так, для $m=3$ компоненты $\beta_2(1)$, $\beta_2(2)$ и $\beta_1(2)$ могут принимать произвольные значения, определяя одну из возможных последовательностей SS1, SS2, ..., SS8 (табл. 3) [4].

Таблица 3

Последовательности Соболя для $m=3$

$m=3$	SS1	SS2	SS3	SS4	SS5	SS6	SS7	SS8
1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0
$\beta_2(1)$ 1 0	0 1 0	1 1 0	0 1 0	0 1 0	1 1 0	1 1 0	0 1 0	1 1 0
$\beta_2(2)$ $\beta_1(2)$ 1	0 0 1	0 0 1	0 1 1	1 0 1	0 1 1	1 0 1	1 1 1	1 1 1
$A(0)$	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0
$A(1)=A(0)+v_0$	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0	1 0 0
$A(2)=A(1)+v_1$	1 1 0	0 1 0	1 1 0	1 1 0	0 1 0	0 1 0	1 1 0	0 1 0
$A(3)=A(2)+v_0$	0 1 0	1 1 0	0 1 0	0 1 0	1 1 0	1 1 0	0 1 0	1 1 0
$A(4)=A(3)+v_2$	0 1 1	1 1 1	0 0 1	1 1 1	1 0 1	0 1 1	1 0 1	0 0 1
$A(5)=A(4)+v_0$	1 1 1	0 1 1	1 0 1	0 1 1	0 0 1	1 1 1	0 0 1	1 0 1
$A(6)=A(5)+v_1$	1 0 1	1 0 1	1 1 1	0 0 1	1 1 1	0 0 1	0 1 1	0 1 1
$A(7)=A(6)+v_0$	0 0 1	0 0 1	0 1 1	1 0 1	0 1 1	1 0 1	1 1 1	1 1 1

Математическая модель, описанная соотношением (2) и матрицей направляющих чисел в виде нижней треугольной матрицы с единичной диагональю (см. табл. 2), может быть расширена для случая адресных последовательностей, относящихся не только к квазислучайным последовательностям. В общем случае в качестве порождающей матрицы направляющих чисел V может быть использована любая квадратная матрица

$$V = \begin{pmatrix} \beta_{m-1}(0) & \beta_{m-2}(0) & \beta_{m-3}(0) & \dots & \beta_0(0) \\ \beta_{m-1}(1) & \beta_{m-2}(1) & \beta_{m-3}(1) & \dots & \beta_0(1) \\ \beta_{m-1}(2) & \beta_{m-2}(2) & \beta_{m-3}(2) & \dots & \beta_0(2) \\ \dots & \dots & \dots & \dots & \dots \\ \beta_{m-1}(m-1) & \beta_{m-2}(m-1) & \beta_{m-3}(m-1) & \dots & \beta_0(m-1) \end{pmatrix}, \quad (3)$$

построенная из m линейно независимых двоичных векторов $v_i = \beta_{m-1}(i)\beta_{m-2}(i)\dots\beta_0(i)$, $i = \overline{0, m-1}$.

Для оценки свойств последовательности Соболя $A(n) = a_{m-1}a_{m-2}a_{m-3}\dots a_1a_0$, используемой в качестве адресной последовательности, в работе [4] введена метрика $M(j)$, $j \in \{0, 1, 2, \dots, m-1\}$, определяющая количество переключений (изменений) j -го разряда a_j кода последовательности

$A(n)$. В общем случае для произвольного значения j величина данной метрики определяется согласно выражению

$$M(j) = \sum_{i=0}^{m-1} \beta_j(i) \times 2^{m-1-i}. \quad (4)$$

Для случая адресных последовательностей Соболя в силу наличия ограничений на значения направляющих чисел выражение для данной метрики принимает вид [4]

$$M(j) = 2^j + \sum_{l=1}^j \beta_j(m-1-j+l) \times 2^{j-l}.$$

Из приведенного соотношения, например, следует, что младший разряд a_0 кода $A(n) = a_{m-1} a_{m-2} a_{m-3} \dots a_1 a_0$ элементов последовательности Соболя изменит свое значение только один раз, так как согласно (4) $M(0) = 1$. Величина $M(1) = 2^1 + \beta_1(m-1) \times 2^0$ в зависимости от значения $\beta_1(m-1)$ может принимать одно из двух значений: 2, если $\beta_1(m-1) = 0$, либо 3, если $\beta_1(m-1) = 1$. Это означает, что следующий разряд a_1 кода $x_n = a_{m-1} a_{m-2} a_{m-3} \dots a_1 a_0$ вне зависимости от длины последовательности Соболя изменит свое значение два либо три раза. В случае старшего разряда a_{m-1} имеем $M(m-1) = 2^{m-1} + \beta_{m-1}(1) \times 2^{m-2} + \beta_{m-1}(2) \times 2^{m-3} + \dots + \beta_{m-1}(m-1) \times 2^0$. Минимальное значение $M_{\min}(m-1)$ количества переключений для старшего разряда равняется 2^{m-1} в случае, когда $\beta_{m-1}(1) = \beta_{m-1}(2) = \dots = \beta_{m-1}(m-1) = 0$, а максимальное значение $M_{\max}(m-1)$ равняется $2^m - 1$ в случае, когда $\beta_{m-1}(1) = \beta_{m-1}(2) = \dots = \beta_{m-1}(m-1) = 1$ [4].

Для произвольного разряда $j \in \{0, 1, 2, \dots, m-1\}$ имеем $M_{\min}(j) = 2^j$, а $M_{\max}(j) = 2^{j+1} - 1$, откуда следует, что всегда $M(j+1) - M(j) > 0$ для произвольных значений v_i , представляющих собой нижнюю треугольную матрицу с единичной диагональю (см. табл. 2). Таким образом, количество переключений возрастает с ростом индекса j разряда кода $A(n) = a_{m-1} a_{m-2} a_{m-3} \dots a_1 a_0$ элементов последовательности Соболя. Приведенные оценки подтверждаются для частного случая последовательностей SS1, SS2, ..., SS8 для $m = 3$, представленных в табл. 3.

В качестве меры различия адресных последовательностей в [13] предложено использование среднего расстояния Хэмминга $AHD[A(n), A(n+k)]$ между двумя последовательностями адресов $A(n)$ и $A(n+k)$, которое для случая $k = 1$ вычисляется согласно выражению

$$AHD[A(n), A(n+1)] = \frac{SHD[A(n), A(n+1)]}{2^m - 1}. \quad (5)$$

Здесь $SHD[A(n), A(n+1)]$ представляет собой суммарное расстояние Хэмминга, которое равняется общему количеству переключений разрядов кода $A(n) = a_{m-1} a_{m-2} a_{m-3} \dots a_1 a_0$ при переходе к адресу $A(n+1)$. Для случая последовательностей Соболя максимальное $AHD_{\max}[A(n), A(n+1)]$ и минимальное $AHD_{\min}[A(n), A(n+1)]$ значения данной характеристики определяются согласно

$$\begin{aligned} AHD_{\max}[A(n), A(n+1)] &= \frac{\sum_{j=0}^{m-1} M_{\max}(j)}{2^m - 1} = \frac{\sum_{j=0}^{m-1} 2^{j+1} - 1}{2^m - 1} = 2 - \frac{m}{2^m - 1}; \\ AHD_{\min}[A(n), A(n+1)] &= \frac{\sum_{j=0}^{m-1} M_{\min}(j)}{2^m - 1} = \frac{\sum_{j=0}^{m-1} 2^j}{2^m - 1} = 1. \end{aligned} \quad (6)$$

Значения элементов нижней треугольной матрицы (см. табл. 2) модифицированных направляющих чисел задают определенный вид последовательности Соболя. Например, в случае когда $\beta_{m-1-i}(i) = 0$ для всех $j < i$, последовательность Соболя представляет собой одну из по-

следовательностей кода Грея [4]. Для случая когда $\beta_{m-1-i}(i) = 1$ для всех $j < i$, последовательность Соболя (см. табл. 3) представляет собой последовательность Корпута [2, 4].

Приведенный анализ позволяет сформулировать следующее утверждение.

Утверждение 1. Модифицированная последовательность Соболя определяется видом порождающей матрицы V , представляющей собой нижнюю треугольную матрицу (унитреугольную матрицу) с единичной главной диагональю.

Характерной особенностью последовательностей Соболя является возрастающая переключательная активность $M(j)$ разрядов кода $a_{m-1}a_{m-2}a_{m-3}\dots a_1a_0$ при возрастании их индексов. Для подобных последовательностей всегда $M(j + 1) - M(j) > 0$.

3. Адресные последовательности

Обобщенная математическая модель, представленная в предыдущем разделе, является расширением математической модели, используемой для модифицированных последовательностей Соболя [2, 4]. Основой данной модели является вид порождающей матрицы V (3), которая и определяет основные свойства адресных последовательностей и идентифицирует их подмножества. Рассмотрим некоторые из них.

1. *Пересчетные адресные последовательности.* В случае пересчетных (счетчиковых) последовательностей, формируемых двоичными пересчетными схемами (счетчиками), наблюдается обратная зависимость для переключательной активности $M(j)$ по сравнению с последовательностями Соболя: частота изменений старшего разряда кода $a_{m-1}a_{m-2}a_{m-3}\dots a_1a_0$ является минимальной, а количество переключений в конкретном разряде уменьшается с ростом его индекса j .

Для получения пересчетных последовательностей необходимо сформировать порождающую матрицу в соответствии с утверждением 2.

Утверждение 2. Пересчетная последовательность определяется видом порождающей матрицы V , представляющей собой нижнюю треугольную матрицу относительно побочной единичной диагонали.

Для произвольной пересчетной последовательности элементы матрицы V имеют определенные значения $\beta_i(i) = 1$ и $\beta_j(i) = 0$ для $j > i$ и произвольные значения $\beta_j(i) \in \{0, 1\}$ для $j < i$, $i, j \in \{0, 1, 2, \dots, m - 1\}$. Примеры подобных последовательностей $CS1, CS2, \dots, CS8$ для $m = 3$ приведены в табл. 4.

Таблица 4

Пересчетные последовательности для $m = 3$

$m = 3$	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8
0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1
0 1 $\beta_0(1)$	0 1 0	0 1 1	0 1 0	0 1 0	0 1 1	0 1 1	0 1 0	0 1 1
1 $\beta_1(2)$ $\beta_0(2)$	1 0 0	1 0 0	1 1 1	1 0 1	1 1 0	1 0 1	1 1 0	1 1 1
$A(0)$	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0
$A(1) = A(0) + v_0$	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1	0 0 1
$A(2) = A(1) + v_1$	0 1 1	0 1 0	0 1 1	0 1 1	0 1 0	0 1 0	0 1 1	0 1 0
$A(3) = A(2) + v_0$	0 1 0	0 1 1	0 1 0	0 1 0	0 1 1	0 1 1	0 1 0	0 1 1
$A(4) = A(3) + v_2$	1 1 0	1 1 1	1 0 1	1 1 1	1 0 1	1 1 0	1 0 0	1 0 0
$A(5) = A(4) + v_0$	1 1 1	1 1 0	1 0 0	1 1 0	1 0 0	1 1 1	1 0 1	1 0 1
$A(6) = A(5) + v_1$	1 0 1	1 0 1	1 1 0	1 0 0	1 1 1	1 0 0	1 1 1	1 1 0
$A(7) = A(6) + v_0$	1 0 0	1 0 0	1 1 1	1 0 1	1 1 0	1 0 1	1 1 0	1 1 1

Из табл. 4 видно, что $CS8$ представляет собой стандартную двоичную последовательность (Counter Sequence), т. е. последовательность целых чисел, представленных в двоичной системе счисления [12, 13]. Последовательность $CS1$ является классическим отраженным кодом Грея (Reflected Gray Code) [22]. Обобщающим свойством рассматриваемого подмножества ад-

ресных последовательностей является убывающая переключаемая активность разрядов кода $a_{m-1}a_{m-2}a_{m-3}\dots a_1a_0$ адреса при возрастании их индексов.

Общее количество пересчетных последовательностей, так же как и последовательностей Соболя, формируемых согласно (2), определяется величиной $q = 2^{(m^2-m)/2}$.

Для оценки свойств пересчетной адресной последовательности используем метрику $M(j)$ (4), которая для данного случая принимает вид

$$M(j) = 2^{m-1-j} + \sum_{l=1}^{m-1-j} \beta_j(j+l) \times 2^{m-1-j-l}.$$

Полученные значения метрики $M(j)$ позволяют оценить среднюю переключаемую активность для генерируемой адресной последовательности $A(n)$, которая определяется средним расстоянием Хэмминга. Так, максимально возможная переключаемая активность для последовательности Соболя и пересчетной последовательности достигается при всех единичных значениях $\beta_j(i)$ для $j < i$, участвующих в соотношениях для $M(j)$. Эта величина, так же как и минимальная переключаемая активность, аналогично последовательностям Соболя определяется соотношениями $M_{max}(j) = 2^{j+1} - 1$ и $M_{min}(j) = 2^j$. Соответственно среднее расстояние Хэмминга $AHD[A(n), A(n+1)]$ между двумя последовательными адресами пересчетной адресной последовательности вычисляется согласно соотношению (6).

2. *Модифицированные квазислучайные и пересчетные последовательности.* Анализ матрицы (3) позволяет идентифицировать еще два множества адресных последовательностей, для которых максимально возможная переключаемая активность принимает заметно большее значение, чем для оригинальных модифицированных последовательностей Соболя и пересчетных последовательностей. Эти последовательности назовем модифицированными последовательностями Соболя и пересчетными последовательностями с максимальной переключаемой активностью, соответствующими двум следующим утверждениям.

Утверждение 3. Порождающая матрица V модифицированной последовательности Соболя с максимальной переключаемой активностью представляет собой верхнюю треугольную матрицу с единичной побочной диагональю.

Утверждение 4. Порождающая матрица V пересчетной последовательности с максимальной переключаемой активностью представляет собой верхнюю треугольную матрицу относительно единичной главной диагонали.

Примеры последовательностей из двух указанных множеств приведены в табл. 5.

Таблица 5

Модифицированные квазислучайные и пересчетные последовательности для $m = 3$

$m = 3$	SS_{m1}	SS_{m2}	SS_{m3}	SS_{m4}	$m = 3$	CS_{m1}	CS_{m2}	CS_{m3}	CS_{m4}
$\beta_2(0)\beta_1(0) 1$	1 1 1	1 1 1	1 0 1	0 0 1	$1 \beta_1(0) \beta_0(0)$	1 1 1	1 1 1	1 0 1	1 0 0
$\beta_2(1) 1 \ 0$	1 1 0	0 1 0	1 1 0	0 1 0	$0 \ 1 \ \beta_0(1)$	0 1 1	0 1 0	0 1 1	0 1 0
$1 \ 0 \ 0$	1 0 0	1 0 0	1 0 0	1 0 0	$0 \ 0 \ 1$	0 0 1	0 0 1	0 0 1	0 0 1
$A(0)$	0 0 0	0 0 0	0 0 0	0 0 0	$A(0)$	0 0 0	0 0 0	0 0 0	0 0 0
$A(1) = A(0) + v_0$	1 1 1	1 1 1	1 0 1	0 0 1	$A(1) = A(0) + v_0$	1 1 1	1 1 1	1 0 1	1 0 0
$A(2) = A(1) + v_1$	0 0 1	1 0 1	0 1 1	0 1 1	$A(2) = A(1) + v_1$	1 0 0	1 0 1	1 1 0	1 1 0
$A(3) = A(2) + v_0$	1 1 0	0 1 0	1 1 0	0 1 0	$A(3) = A(2) + v_0$	0 1 1	0 1 0	0 1 1	0 1 0
$A(4) = A(3) + v_2$	0 1 0	1 1 0	0 1 0	1 1 0	$A(4) = A(3) + v_2$	0 1 0	0 1 1	0 1 0	0 1 1
$A(5) = A(4) + v_0$	1 0 1	0 0 1	1 1 1	1 1 1	$A(5) = A(4) + v_0$	1 0 1	1 0 0	1 1 1	1 1 1
$A(6) = A(5) + v_1$	0 1 1	0 1 1	0 0 1	1 0 1	$A(6) = A(5) + v_1$	1 1 0	1 1 0	1 0 0	1 0 1
$A(7) = A(6) + v_0$	1 0 0	1 0 0	1 0 0	1 0 0	$A(7) = A(6) + v_0$	0 0 1	0 0 1	0 0 1	0 0 1

Для модифицированной последовательности Соболя с максимальной переключаемой активностью значение $M(j)$ (4) вычисляется согласно выражению

$$M(j) = 2^{m-1-j} + \sum_{l=1}^j \beta_j(j-l) \times 2^{m-1-j+l}.$$

Аналогичное выражение справедливо и для пересчетной последовательности с максимальной переключающей активностью:

$$M(j) = 2^j + \sum_{l=1}^{m-1-j} \beta_j(m-1-j-l) \times 2^{j+l}.$$

Как в случае модифицированной последовательности Соболя с максимальной активностью (см. утверждение 3), так и пересчетной последовательности с максимальной переключающей активностью (утверждение 4), используя метрику $M(j)$, можно получить аналогичные соотношения для среднего расстояния Хэмминга $AHD[A(n), A(n+1)]$. В этом случае $AHD_{min}[A(n), A(n+1)]$ также равняется 1, что соответствует последовательностям кода Грея $SS_m4 = CS1$ и $CS_m4 = SS1$ (см. табл. 3 и 4). В то же время максимальная переключающая активность для обеих последовательностей SS_m и CS_m определяется согласно соотношению

$$AHD_{max}[A(n), A(n+1)] = m-1 + \frac{m}{2^m - 1}. \quad (7)$$

Из последнего выражения видно, что максимальное среднее расстояние Хэмминга (7) значительно больше аналогичной величины для оригинальных модифицированных последовательностей Соболя и пересчетных последовательностей (6), практически на величину, равную $m-3$.

Общее количество модифицированных последовательностей Соболя и пересчетных последовательностей с максимальной переключающей активностью определяется величиной $q = 2^{(m^2-m)/2}$.

3. *Адресные последовательности с минимальной переключающей активностью.* К данному множеству относятся упоминаемые ранее последовательности кода Грея, пример которых присутствует во всех рассмотренных множествах адресных последовательностей. Действительно, $SS1$ принадлежит множеству модифицированных последовательностей Соболя, SS_m4 относится к множеству модифицированных последовательностей Соболя с максимальной переключающей активностью, $CS1 = SS_m4$ представляет собой пересчетную последовательность, $CS_m4 = SS1$ соответствует пересчетной последовательности с максимальной переключающей активностью. Каждой из указанных адресных последовательностей присущи свойства множеств последовательностей, к которым они принадлежат.

В общем случае адресная последовательность с минимальной переключающей активностью (минимальным расстоянием Хэмминга), формируемая согласно (2), обеспечивается матрицей V (3) с минимальным количеством единичных значений. Для произвольного случая подобная матрица строится в соответствии со следующим утверждением.

Утверждение 5. Порождающая матрица V для адресной последовательности с минимальной переключающей активностью состоит из m отличающихся друг от друга строк, каждая из которых содержит по одной единице.

Согласно приведенному утверждению подобная порождающая матрица характеризуется тем, что она представляет собой множество отличающихся друг от друга столбцов, содержащих по одной единице. Из этого следует, что каждый столбец обеспечивает одно из возможных значений ($2^0, 2^1, 2^2, \dots, 2^{m-1}$) переключающей активности. В силу этого последовательность столбцов в матрице, удовлетворяющей утверждению 5, не влияет на среднее значение переключающей активности. Для последовательностей $A(n)$, формируемых с использованием подобных порождающих матриц V (3), $AHD[A(n), A(n+1)] = 1$, а общее их количество равняется $m!$.

4. *Адресные последовательности с высокой переключающей активностью.* С целью получения последовательностей адресов с максимальным расстоянием Хэмминга между соседними адресами $A(n)$ и $A(n + 1)$ за основу может быть взята последовательность переключений T_{m-1} отраженного кода Грея [7, 11]. Такая последовательность имеет инверсную интерпретацию, т. е. текущий элемент последовательности T_{m-1} определяет индекс неизменяемого разряда $A(n)$ при переходе к следующему адресу $A(n + 1)$. В этом случае все разряды, кроме одного, при переходе от $A(n)$ к $A(n + 1)$ будут изменяться. Соответственно расстояние Хэмминга между двумя последовательными адресами будет равняться $m - 1$. Такие последовательности получили название последовательностей анти-Грея [23].

Для обеспечения равенства расстояния Хэмминга величине $m - 1$ в рамках выбранной математической модели генерирования адресных последовательностей очевидным является использование матрицы направляющих чисел (3), инверсной по отношению к случаю, соответствующему утверждению 5. Отметим, что данное утверждение позволяет генерировать адресные последовательности с минимальным расстоянием Хэмминга, равным 1. Рассмотрим несколько примеров использования матриц (3), инверсных по отношению к аналогичным матрицам для последовательностей кода Грея.

Таблица 6

Примеры генерирования последовательностей с максимальной переключающей активностью

	$m = 2$	$m = 3$	$m = 4$		$m = 4$
T_{m-1}	$V = \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$	$V = \begin{matrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{matrix}$	$V = \begin{matrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$	T_{m-1}	$V = \begin{matrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{matrix}$
	0 0	0 0 0	0 0 0 0	3	1 1 0 0
0	1 0	1 1 0	1 1 1 0	0	0 0 1 0
1	1 1	0 1 1	0 0 1 1	1	1 1 1 1
0	0 1	1 0 1	1 1 0 1	0	0 0 0 1
2		1 1 0	0 1 1 0	2	1 0 1 0
0		0 0 0	1 0 0 0	0	0 1 0 0
1		1 1 0	0 1 0 1	1	1 0 0 1
0		0 1 1	1 0 1 1	0	0 1 1 1

В табл. 6 сформированы все возможные комбинации адресов из двух ($m = 2$) и четырех ($m = 4$) бит и, таким образом, получены последовательности анти-Грея с расстоянием Хэмминга, равным $m - 1$. В то же время для $m = 3$ данный подход оказался неработоспособным в силу линейной зависимости между столбцами порождающей матрицы V для нечетных значений m [23].

Адресная последовательность с переключающей активностью, равной $m - 1$ для четных значений m , формируемая согласно соотношению (2), обеспечивается матрицей V (3), удовлетворяющей следующему утверждению.

Утверждение 6. Порождающая матрица V для адресной последовательности с переключающей активностью, равной $m - 1$ для четных значений m , состоит из m отличающихся друг от друга строк, каждая из которых содержит $m - 1$ единиц.

Для последовательностей, формируемых с использованием подобных порождающих матриц V (3), $AHD[A(n), A(n+1)] = m - 1$. Подобные последовательности формируются только для четных значений m , а общее их количество равняется $m!$.

5. *Последовательности с предельно возможной переключающей активностью.* В качестве операции, позволяющей достичь максимальной переключающей активности при генерировании адресных последовательностей, весьма эффективным является использование операции инвертирования всех разрядов предыдущего адреса $A(n)$ для получения последующего его значения $A(n + 1)$. В этом случае расстояние Хэмминга равняется предельно возможному

значению, равному m . В терминах математической модели, используемой в данной статье, инвертирование всех разрядов кода адреса эквивалентно применению одного из направляющих чисел v_i , представляющих собой единичный вектор, т. е. $v_i = 111\dots 1$. Очевидно, что для достижения максимального среднего расстояния Хэмминга необходимо, чтобы таким направляющим числом было число v_0 . Последующее инвертирование адреса $A(n + 1)$ для получения $A(n + 2)$ приведет к повторению исходного адреса, т. е. $A(n + 2) = A(n)$. Для целей обеспечения предельно возможной средней переключающей активности необходимо перемежающее получение расстояний Хэмминга, равных m и $m - 1$ для последовательных пар адресов $A(n)$, $A(n + 1)$ и $A(n + 1)$, $A(n + 2)$ [23]. Обобщая приведенные рассуждения, можно сформулировать следующее утверждение.

Утверждение 7. Порождающая матрица V для адресной последовательности с предельной максимальной переключающей активностью состоит из одного единичного столбца и $m - 1$ отличающихся друг от друга столбцов, содержащих по одному нулевому значению в каждой из $m - 1$ строк, кроме первой.

Примеры последовательностей из множества, определяемого утверждением 7, для $m = 3$ приведены в табл. 7.

Таблица 7

Последовательности с предельной максимальной переключающей активностью для $m = 3$

$m = 3$	LS1	LS2	LS3	LS4	LS5	LS6
V	1 1 1 0 1 1 1 0 1	1 1 1 1 0 1 0 1 1	1 1 1 0 1 1 1 1 0	1 1 1 1 1 0 0 1 1	1 1 1 1 1 0 1 0 1	1 1 1 1 0 1 1 1 0
$A(0)$	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0	0 0 0
$A(1) = A(0) + v_0$	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1	1 1 1
$A(2) = A(1) + v_1$	1 0 0	0 1 0	1 0 0	0 0 1	0 0 1	0 1 0
$A(3) = A(2) + v_0$	0 1 1	1 0 1	0 1 1	1 1 0	1 1 0	1 0 1
$A(4) = A(3) + v_2$	1 1 0	1 1 0	1 0 1	1 0 1	0 1 1	0 1 1
$A(5) = A(4) + v_0$	0 0 1	0 0 1	0 1 0	0 1 0	1 0 0	1 0 0
$A(6) = A(5) + v_1$	0 1 0	1 0 0	0 0 1	1 0 0	0 1 0	0 0 1
$A(7) = A(6) + v_0$	1 0 1	0 1 1	1 1 0	0 1 1	1 0 1	1 1 0

Переключающая активность рассмотренной последовательности состоит из суммы переключающих активностей разрядов адресов, определяемых утверждением 7. Для любой подобной матрицы эта сумма включает слагаемые $(2^m - 1) - 0$, $(2^m - 1) - 1$, $(2^m - 1) - 2^1$, $(2^m - 1) - 2^2$, ..., $(2^m - 1) - 2^{m-2}$. Для матриц V , отвечающих данным требованиям, получим

$$AHD_{\max}[A(n), A(n + 1)] = m - \frac{1}{2} + \frac{1}{2(2^m - 1)}. \tag{8}$$

Очевидно, что данное значение $AHD_{\max}[A(n - 1), A(n)]$ (8) является максимально достижимым [7, 11, 23].

4. Адресные последовательности для многократного тестирования ОЗУ

Тестирование современных ОЗУ с использованием многократных маршевых тестов и увеличение их эффективности за счет применения различных сочетаний адресных последовательностей является весьма актуальной задачей [3, 5, 7]. Очевидно, что для достижения максимальной полноты покрытия неисправностей ОЗУ в результате многократного применения маршевых тестов необходимо, чтобы адресные последовательности, применяемые для каждого прохода теста, максимально отличались друг от друга. В простейшем случае это означает, что на одной и той же позиции двух последовательностей должны быть адреса, максимально отличные друг от друга. Подобное требование реализовывалось в классических маршевых тестах [9]. В со-

пряженных фазах теста, как правило, применяется возрастающая $A_{incr}(n) = 000\dots00, 000\dots01, 000\dots10, \dots, 111\dots11$ (в десятичной системе счисления: $0, 1, 2, \dots, 2^m - 1$) и убывающая последовательности адресов $A_{decr}(n) = 111\dots11, 111\dots10, 111\dots01, \dots, 000\dots00$ ($2^m - 1, 2^m - 2, 2^m - 3, \dots, 0$). Доказано, что данное сочетание адресных последовательностей является оптимальным с точки зрения покрывающей способности неисправностей ОЗУ [7, 9]. В общем случае выбор набора адресных последовательностей требует обоснованного формального подхода. В качестве меры различия адресных последовательностей часто используется арифметическое расстояние [7, 12].

Классическое расстояние Минковского между двумя адресными последовательностями $A_v(n)$ и $A_w(n)$, $n \in \{0, 1, 2, \dots, 2^m - 1\}$, при выполнении равенства $\lambda = 1$ для его параметра λ принимает вид арифметического расстояния [24]:

$$D_{Manh.}[A_v(n), A_w(n)] = \sum_{n=0}^{2^m-1} |A_v(n) - A_w(n)|. \quad (9)$$

Расстояние Минковского, частным случаем которого является арифметическое расстояние (9), как правило, используется как мера различия, а не мера подобия между объектами [25]. Для данной меры различия в случае двукратного тестирования ОЗУ, когда последовательно применяются адресные последовательности $A_{incr}(n)$ и $A_{decr}(n)$, значение метрики различия оказывается максимальным и определяется согласно выражению [12]

$$D_{Manh_max}[A_{incr}(n), A_{decr}(n)] = \sum_{n=0}^{2^m-1} |2^m - 2n - 1| = 2^{2m-1}.$$

Весьма эффективным подходом для получения различных адресных последовательностей в терминах арифметического расстояния является инвертирование определенных разрядов $A_{incr}(n)$ для получения модифицированной последовательности $A_{mod}(n)$ из исходной $A_{incr}(n)$ [12].

Проведенный анализ свидетельствует о необходимости генерирования адресов $A_{decr}(n)$ в обратной последовательности по отношению к оригинальной адресной последовательности $A_{incr}(n)$ и различных их модификаций как результата инвертирования определенных разрядов $A_{incr}(n)$. В рамках выбранной модели генерирования адресных последовательностей (2) в силу эквивалентности операций сложения и вычитая по модулю два (\oplus), а также симметричности последовательности переключений T_{m-1} отраженного кода Грея формирование последовательности убывающих $A_{decr}(n)$ адресов по отношению к $A_{incr}(n)$ соответствует следующему утверждению.

Утверждение 8. Убывающая последовательность адресов $A_{decr}(n)$, $n \in \{0, 1, 2, \dots, 2^m - 1\}$, по отношению к возрастающей последовательности $A_{incr}(n)$, для которой $A_{incr}(i) = A_{decr}(2^m - 1 - i)$, генерируется с использованием соотношения (2) и той же порождающей матрицы V , что и для генерирования $A_{incr}(n)$ при начальном адресе $A_{decr}(0)$, равном $A_{incr}(2^m - 1)$.

В силу эквивалентности операции сложения по модулю два с единицей ($a_i \oplus 1$) и операции отрицания $\overline{a_i}$ справедливо следующее утверждение.

Утверждение 9. Модифицированная последовательность адресов $A_{mod}(n)$, $n \in \{0, 1, 2, \dots, 2^m - 1\}$ как результат инвертирования определенных разрядов возрастающей последовательности $A_{incr}(n)$ генерируется с использованием соотношения (2) и той же порождающей матрицы V , что и для формирования $A_{incr}(n)$, при начальном адресе $A_{incr}(0)$, содержащем в инвертируемых разрядах единичные значения, а в остальных разрядах – нулевые значения.

Отметим, что данные утверждения справедливы для последовательностей, описанных в разд. 3 и сформированных на основании соотношения (2) и порождающей матрицы V (3). В качестве примера приведем последовательность $LS1$ с предельной максимальной переключательной активностью для $m = 3$ (см. табл. 7). Рассматривая $LS1$ как возрастающую последовательность LS_{incr} , в табл. 8 приведем ряд ее модификаций.

Таблица 8

Примеры модификаций последовательности $LS1$

$m = 3$	$LS_{incr} = a_2 a_1 a_0$	$LS_{decr} = a_2 a_1 a_0$	$LS_{mod1} = \bar{a}_2 a_1 a_0$	$LS_{mod2} = a_2 a_1 \bar{a}_0$
V	1 1 1 0 1 1 1 0 1	1 1 1 0 1 1 1 0 1	1 1 1 0 1 1 1 0 1	1 1 1 0 1 1 1 0 1
$A(0)$	0 0 0	1 0 1	1 0 0	0 0 1
$A(1) = A(0) + v_0$	1 1 1	0 1 0	0 1 1	1 1 0
$A(2) = A(1) + v_1$	1 0 0	0 0 1	0 0 0	1 0 1
$A(3) = A(2) + v_0$	0 1 1	1 1 0	1 1 1	0 1 0
$A(4) = A(3) + v_2$	1 1 0	0 1 1	0 1 0	1 1 1
$A(5) = A(4) + v_0$	0 0 1	1 0 0	1 0 1	0 0 0
$A(6) = A(5) + v_1$	0 1 0	1 1 1	1 1 0	0 1 1
$A(7) = A(6) + v_0$	1 0 1	0 0 0	0 0 1	1 0 0

Действительно, с помощью начального адреса $LS_{decr}(0) = LS_{incr}(2^3 - 1) = 1\ 0\ 1$ для порождающей матрицы, используемой для LS_{incr} , формируется последовательность LS_{decr} , обратная по отношению к LS_{incr} (см. табл. 8), что соответствует утверждению 8. В то же время LS_{mod1} и LS_{mod2} , приведенные в табл. 8, соответствуют утверждению 9.

Заключение

В статье рассмотрена модель генерирования адресных последовательностей как результат обобщения экономичного способа Антонова и Салеева, используемого для формирования последовательностей Соболя. Применение предложенной модели позволяет формировать широкий спектр адресных последовательностей – от простейших квазислучайных последовательностей до последовательностей с различными свойствами, необходимыми для тестирования ОЗУ. В рамках предложенной модели оказывается возможным формирование множеств адресных последовательностей, имеющих большие значения арифметических расстояний между ними.

Список литературы

1. Иванюк, А.А. Проектирование встраиваемых цифровых устройств и систем / А.А. Иванюк. – Минск : Бестпринт, 2012. – 338 с.
2. Ярмолик, С.В. Квазислучайное тестирование вычислительных систем / С.В. Ярмолик, В.Н. Ярмолик // Информатика. – 2013. – № 3 (39). – С. 92–103.
3. Semiconductor Intellectual Property: Continuing on the Path to Growth. ASIC IP report. Semico Research Corp. – Seville, Spain, 2007. – 100 p.
4. Yarmolik, V.N. Generating Modified Sobol Sequences for Multiple Run March Memory Tests / V.N. Yarmolik, S.V. Yarmolik // Automatic Control and Computer Sciences. – 2013. – Vol. 47, № 5. – P. 242–247.
5. VLSI Test Principles and Architectures Design for Testability / L.T. Wang [et al.]. – San Francisco : Morgan Kaufmann Publishers is an imprint of Elsevier, 2006. – 808 p.
6. Sharma, A.K. Semiconductor Memories: Technology, Testing, and Reliability / A.K. Sharma. – London : John Wiley & Sons, 2002. – 480 p.
7. Ярмолик, С.В. Маршевые тесты для тестирования ОЗУ / С.В. Ярмолик, А.П. Занкович, А.А. Иванюк. – Минск : Издательский центр БГУ, 2009. – 270 с.
8. Sharma, A.K. Advanced Semiconductor Memories: Architectures, Designs, and Applications / A.K. Sharma. – London : John Wiley & Sons, 2003. – 652 p.
9. Goor, A.J. Testing Semiconductor Memories, Theory and Practice / A.J. Goor. – UK, Chichester : John & Sons, 1991. – 536 p.
10. Nicolaidis, M. Theory of Transparent BIST for RAMs / M. Nicolaidis // IEEE Transaction on Computers. – 1996. – Vol. 45, № 10. – P. 1141–1156.

11. Yarmolik, S.V. Address sequences and backgrounds with different Hamming distance for multiple run March tests / S.V. Yarmolik // IEEE Int. J. Appl. Comput. Sci. – 2008. – Vol. 18, № 3. – P. 329–339.
12. Ярмолик, С.В. Многократные неразрушающие маршевые тесты с изменяемыми адресными последовательностями / С.В. Ярмолик, В.Н. Ярмолик // Автоматика и телемеханика. – 2007. – № 4. – С. 126–137.
13. Ярмолик, В.Н. Адресные последовательности для многократных маршевых тестов / В.Н. Ярмолик, С.В. Ярмолик // Автоматика и вычислительная техника. – 2006. – № 5. – С. 59–68.
14. Halton, J.H. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals / J.H. Halton // Numerische Mathematics. – 1960. – Vol. 2, № 1. – P. 84–90.
15. Niederreiter, H. Point sets and sequences with small discrepancy / H. Niederreiter // Monatshefte fur Mathematik. – 1987. – Vol. 104, № 4. – P. 273–337.
16. Chi, H. Computational investigations of quasi-random sequences in generating test cases for specification-based tests / H. Chi, E.I. Jones // Proc. of the 38 Winter Conference, ser. WSC'06. Winter Simulation Conf. Monterey. – CA, USA, 2006. – P. 975–980.
17. Соболев, И.М. Точки, равномерно заполняющие многомерный куб / И.М. Соболев. – М. : Знание, 1985. – 32 с.
18. Ярмолик, С.В. Анализ количественных характеристик различия при тестировании ОЗУ / С.В. Ярмолик, А.Н. Курбацкий, В.Н. Ярмолик // Информатика. – 2008. – № 3(19). – С. 90–98.
19. Chen, T.Y. Quasi-Random Testing / T.Y. Chen, R. Merkel // IEEE Transaction on Reliability. – 2007. – Vol. 56, № 3. – P. 562–568.
20. Joe, S. Constructing Sobol sequences with better two-dimensional projections / S. Joe, F. Kuo // SIAM J. Sci. Comput. – 2008. – № 30. – P. 2635–2654.
21. Антонов, И.А. Экономичный способ вычисления ЛПт-последовательностей / И.А. Антонов, В.М. Салеев // Журн. вычисл. матем. и матем. физ. – 1979. – Т. 19, № 1. – С. 243–245.
22. Savage, C. A survey of combinatorial Gray code / C. Savage // SIAM Review. – 1997. – Vol. 39, № 4. – P. 605–629.
23. Yarmolik, S.V. Modified Gray and Counter Sequences for memory test address generation / S.V. Yarmolik, V.N. Yarmolik // Proc. 13th Int. Conf. Mixed Design of Integrated Circuits and Systems (MIXDES'06). – Gdynia, Poland : Technical University of Lodz, 2006. – P. 572–576.
24. Thompson, A.C. Minkowski Geometry / A.C. Thompson. – N.Y. : Cambridge U.P., 1996. – 364 p.
25. Tubbs, J.D. A note on binary template matching / J.D. Tubbs // Pattern Recogn. – 1989. – Vol. 22, № 4. – P. 359–366.

Поступила 22.11.2013

*Белорусский государственный университет
информатики и радиоэлектроники,
Минск, П. Бровки, 6
e-mail: yarmolik10ru@yahoo.com,
yarmolik@cosmostv.by*

V.N. Yarmolik, S.V. Yarmolik

ADDRESS SEQUENCES FOR MULTI RUN RAM TESTING

A universal approach for generation of address sequences with specified properties is proposed and analyzed. A modified version of the Antonov and Saleev algorithm for Sobol sequences generation is chosen as a mathematical description of the proposed method. Within the framework of the proposed universal approach, the Sobol sequences form a subset of the address sequences. Other subsets are also formed, which are Gray sequences, anti-Gray sequences, counter sequences and sequences with specified properties.

ПРАВИЛА ДЛЯ АВТОРОВ

1. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://jinfo.bas-net.by> в формате файлов текстовых редакторов Microsoft Word 97 и Word 2000 для Windows. Основной текст статьи набирается с переносами шрифтом Times New Roman 11 пт, интервал между строками – одинарный, абзацный отступ 1 см, поля по 2,5 см со всех сторон.

2. Объем статьи не должен превышать 12 страниц (включая таблицы, иллюстрации, список литературы), количество иллюстраций – не больше пяти. Допускаются краткие сообщения до трех страниц.

3. Статья должна иметь индекс УДК (универсальная десятичная классификация).

4. Название статьи, фамилии всех авторов и аннотация должны быть переведены на английский язык. Для каждого из авторов приводится развернутое название учреждения с полным почтовым адресом, а также номер телефона и электронный адрес (e-mail) для связи с редакцией.

5. Формулы, иллюстрации, таблицы, встречающиеся в статье, должны быть пронумерованы в соответствии с порядком цитирования в тексте. Ссылки на рисунки и таблицы в тексте обязательны. Необходимо избегать повторения одних и тех же данных в таблицах, графиках и тексте статьи.

Рисунки должны быть представлены в виде файлов формата .cdr, .ai, .wmf, .psd, .jpg, .tif (.tiff) и выполнены с хорошим разрешением в масштабе, позволяющем четко различать надписи и обозначения. Подрисовочные подписи с расшивкой всех позиций, представленных на рисунке, набираются шрифтом гарнитуры основного текста, размер символов 9 пт. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала.

6. Набор формул выполняется в формульных редакторах Microsoft Equation или Math Type и должен быть единообразным по применению шрифтов и знаков по всей статье.

Прямо () набираются: греческие и русские буквы; математические символы (\sin , \lg , ∞); символы химических элементов (C, Cl, CHCl_3); цифры (римские и арабские); векторы; индексы (верхние и нижние), являющиеся сокращениями слов.

Курсивом (~) набираются: латинские буквы – переменные, символы физических величин (в том числе и в индексе).

7. Сокращения в тексте статьи (за исключением единиц измерения) могут быть использованы только после упоминания полного термина. Единицы измерения физических величин следует приводить в Международной системе СИ.

8. Литература приводится автором общим списком в конце статьи. Ссылки на литературу в тексте идут по порядку и обозначаются цифрой в квадратных скобках. Ссылаться на неопубликованные работы не допускается. С примерами оформления библиографического описания в списке литературы можно ознакомиться в приложении 2 к *Инструкции по оформлению диссертации, автореферата и публикаций по теме диссертации* на сайте Высшей аттестационной комиссии Республики Беларусь <http://vak.org.by>.

9. Поступившие в редакцию статьи направляются на рецензирование специалистам. Основным критерием целесообразности публикации является новизна и информативность статьи. Если по рекомендациям рецензента статья возвращается автору на доработку, а переработанная рукопись вновь рассматривается редколлегией, датой поступления считается день получения редакцией ее окончательного варианта. Статьи не по профилю журнала возвращаются авторам после заключения редколлегии.

10. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все вопросы.

11. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантура, соискательство) в год завершения обучения.

12. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее, или статей, принятых к публикации другими изданиями.

13. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи.

Журнал «Информатика» включен Высшей аттестационной комиссией Республики Беларусь в список научных изданий для опубликования результатов диссертационных исследований.

Индексы

00827

для индивидуальных
подписчиков

008272

для предприятий и
организаций