

На правах рукописи

МИРОНЕНКО Мария Андреевна

**Обратные задачи в математическом
моделировании помеховой обстановки на основе
скрытых полумарковских моделей**

**05.13.18 - Математическое моделирование, численные методы и
комплексы программ**

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико-математических наук

Ростов-на-Дону
2019

Работа выполнена на кафедре алгебры и дискретной математики института математики, механики и компьютерных наук имени И.И.Воровича Южного федерального университета.

НАУЧНЫЙ РУКОВОДИТЕЛЬ:

кандидат физико-математических наук, доцент Деундяк Владимир Михайлович

ОФИЦИАЛЬНЫЕ ОППОНЕНТЫ:

доктор физико-математических наук, профессор -

кандидат технических наук, доцент

ВЕДУЩАЯ ОРГАНИЗАЦИЯ:

Защита состоится “____” _____ 200__ г. в ____ час. ____ мин. на заседании диссертационного совета Д 212.208.25 Южного федерального университета по адресу: 347928, Ростовская область, г. Таганрог, ул. Чехова 2, ауд. __.

Отзывы на автореферат просьба направлять по адресу: 347928, Ростовская область, г. Таганрог, пер. Некрасовский 44, Технологический институт Южного федерального университета в г. Таганроге, Ученому секретарю диссертационного совета Д 212.208.25 Брюхомицкому Ю.А.

С диссертацией можно ознакомиться в Зональной научной библиотеке ЮФУ по адресу: 344007, Ростовская область, г. Ростов-на-Дону, ул. Пушкинская, 148.

Автореферат разослан “____” _____ 200__ г.

Ученый секретарь
диссертационного совета Д 212.208.25,
к.т.н.

Брюхомицкий Ю.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования.

Для обеспечения высококачественной передачи информации в последние десятилетия широко используются цифровые системы передачи данных. Цифровая передача данных предоставляет больше возможностей в обработке информации по сравнению с аналоговой, цифровые каналы менее подвержены искажению и интерференции, а применение процедур обнаружения и исправления ошибок делают возможной высокую точность сигнала. Для обнаружения и исправления ошибок эффективно используются средства алгебраического помехоустойчивого кодирования. За последние полвека было разработано большое количество алгебраических кодов, исправляющих ошибки, в частности, коды Хэмминга, Боуза-Чоудхури-Хоквингема, Рида-Соломона, Рида-Маллера, циклические, сверточные и др. В настоящее время продолжается разработка новых кодов и алгоритмов декодирования. Важные результаты в этой области получены М. Суданом, В.М. Сидельниковым, С.А. Осмоловским и ... (еще три строчки фамилий, может быть с конкретизацией областей или (и) стран, городов, институтов) другими учеными.

При применении помехоустойчивого кодирования следует учитывать, что различные коды и алгоритмы их декодирования неодинаково хорошо справляются с разными типами ошибок. Поэтому нужно уметь подбирать для каждого конкретного канала наиболее эффективные методы и средства обнаружения и исправления ошибок. Для решения этой задачи необходимо иметь достаточный объем сведений о возможных помехах в канале и о корректирующих свойствах кода по отношению к ошибкам различной структуры. Подбор помехоустойчивого кода к реальному каналу является сложной задачей оптимизации соотношения между затратами на кодирование и обеспечиваемым качеством передачи информации. Проведение реальных экспериментов, как правило, достаточно трудоемко, поэтому эффективным оказывается применение имитационного моделирования помехоустойчивых каналов связи. Основанные на имитационном моделировании системы оценки применимости схем алгебраического помехоустойчивого кодирования позволяют подбирать помехоустойчивые коды по потокам ошибок, порождаемым математическими моделями источников ошибок. Схема информационной системы оценки применимости схем алгебраического помехоустойчивого кодирования была построена в работах Н.С. Могилевской. Предложенная система позволяет решать такие задачи как согласование помехоустойчивого кода с каналом, изучение корректирующих свойств декодеров, исследование быстродействия работы кодера и декодера и многие другие.

Важную роль в работе предложенной информационной системы играют модели источников потоков ошибок, имитирующие внешние воздействия в канале связи. Моделированием источников потоков ошибок занимались такие ученые как Э.Н. Гильберт, Е.О. Эллиот, Б.М. Игельник, В.И. Петрович, Б.Д. Фричман, В.М. Охорзин, В.О. Колпаков, В.Я. Турин, О.В. Попов, Ю.С. Чье и другие. В настоящее время известно множество классических моделей источников ошибок, однако, как правило математическая модель источника ошибок описывает узкий класс каналов связи, поэтому для исследования корректирующей способности кода

по отношению к различным типам ошибок при проведении имитационных экспериментов необходимо использовать несколько моделей источников ошибок. Это затрудняет проведение имитационных экспериментов, поскольку в их процессе приходится тестировать корректирующую способность кодеков для разных моделей. Представляется более удобным построить общую модель источника ошибок канала, которая учитывала бы q -ичные алфавиты и позволила бы моделировать различные случаи помеховой обстановки путем изменения параметров.

Для проведения имитационных экспериментов, направленных на подбор кодека к конкретному каналу связи, важно, чтобы используемая модель канала связи соответствовала реальному каналу связи, для которого подбирается кодек. Поэтому необходимо научиться решать задачу адекватного представления зарегистрированной в реальном канале последовательности ошибок математической моделью источника ошибок, чтобы впоследствии использовать эту модель в качестве генератора потоков ошибок при проведении имитационных экспериментов с целью подбора кодека.

Таким образом, актуальным представляется построение общей математической модели источника ошибок в канале передачи данных и разработка алгоритма подбора адекватной модели источника ошибок для конкретного канала связи.

Целью диссертационной работы является построение общей математической модели источника ошибок в канале передачи данных, разработка алгоритмов, численных методов и программного комплекса, позволяющих осуществлять генерацию потоков ошибок в соответствие с построенной моделью и подбор адекватной математической модели источника ошибок для конкретного канала связи.

Задачами диссертационной работы являются:

1. Провести содержательный анализ и систематизацию сведений о существующих способах моделирования источников потоков ошибок и определить математический аппарат, позволяющий построить такую общую математическую модель источника ошибок, для которой возможно решение задачи подбора конкретной модели, соответствующей зарегистрированной в канале передачи данных последовательности ошибок.
2. Разработать общую модель источника ошибок на основе теории скрытых марковских и полумарковских моделей.
3. Разработать численные методы, позволяющие для реального канала передачи данных подбирать адекватную математическую модель источника ошибок.
4. Усовершенствовать информационную систему оценки применимости схем алгебраического помехоустойчивого кодирования, расширив ее функционал возможностью подбирать подходящую модель источника ошибок для исследуемого реального канала связи.
5. Разработать программный комплекс для исследования моделей источников ошибок, позволяющий как генерировать потоки ошибок с помощью разработанных ранее моделей, так и подбирать наилучшую модель по реальному потоку ошибок в канале.

Методология и методы исследования. При выполнении работы использовались методы алгебры, теории конечных автоматов, теории вероятностей, математической статистики, компьютерного эксперимента. Существенно применялась теория скрытых марковских и полумарковских моделей, в частности, методы решения задач оценивания. Для реализации вычислительного ядра программного комплекса использовался язык программирования C#, пользовательский интерфейс реализован посредством интерфейса программирования приложений Windows Forms. При разработке архитектуры программного комплекса и в ходе программной реализации соблюдались принципы объектно-ориентированного проектирования, исходный код программного комплекса покрыт модульными и интеграционными тестами.

Достоверность научных результатов подтверждается точностью постановок задач, совпадением частных случаев предлагаемых моделей с классическими моделями, обоснованностью принятых предположений и допущений, приведенным доказательством новых и корректным применением известных теоретических положений, корректным применением методов и алгоритмов, устойчивой работой разработанного программного комплекса и результатами численных экспериментов, проведенных на модельных примерах.

Объектом исследования диссертационной работы являются потоки ошибок в цифровых помехоустойчивых каналах связи.

Предметом исследования выступают методы моделирования и анализа потоков ошибок в цифровых каналах передачи данных.

Тематика работы соответствует п.1 «Разработка новых математических методов моделирования объектов и явлений.», п.3 «Разработка, обоснование и тестирование эффективных вычислительных методов с применением современных компьютерных технологий», п.4 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента» и п.8 «Разработка систем компьютерного и имитационного моделирования» паспорта специальности 05.13.18 – «Математическое моделирование, численные методы и комплексы программ».

Научная новизна. В диссертации получены следующие новые научные и практические результаты:

- *в области математического моделирования:*

- разработаны три новых модели источника ошибок в q -ичном цифровом канале передачи данных — скрытая марковская модель источника ошибок, скрытая полумарковская модель Фергюсона и скрытая полумарковская QP -модель, позволяющие имитировать различные случаи помеховой обстановки, совпадающие в частных случаях со многими широко применяющимися моделями источников ошибок и отличающиеся от существующих возможностью решения для них задачи подбора модели с такими параметрами, что она сможет адекватно имитировать конкретный реальный канал связи. Все предложенные модели позволяют моделировать потоки ошибок в q -ичном цифровом канале с несколькими физическими состояниями. Скрытая марковская модель источника

ошибок является обобщением классических моделей Гильберта и Фричмана и предполагает, что длительность пребывания в состоянии описывается геометрическим распределением. Отличие скрытой полумарковской модели Фергюсона заключается в явном задании длительности пребывания модели в состоянии, при этом распределение вероятностей ошибок внутри состояния является равномерным. Наиболее важной с точки зрения научной новизны является скрытая полумарковская QP -модель, позволяющая как и модель Фергюсона явно моделировать длительность пребывания в состоянии, а также задавать произвольное распределение вероятностей ошибок на протяжении состояния.

- построена модификация модели информационной системы оценки применимости схем помехоустойчивого алгебраического кодирования, в рамках которой впервые предложено использование скрытых полумарковских моделей в качестве базовых моделей источника ошибок и добавлен новый модуль подбора адекватной реальному каналу модели источника ошибок (из базы данных моделей источников ошибок информационной системы). Такая модификация позволяет решить в рамках информационной системы оценки применимости схем помехоустойчивого алгебраического кодирования ранее не ставившийся вопрос адекватности имитационной модели канала реальному каналу связи.
- *в области численных методов:*
 - впервые предложен и теоретически обоснован алгоритм подбора по зарегистрированной в реальном канале передачи данных последовательности ошибок адекватной математической модели источника ошибок из заранее определенного набора моделей. Этот алгоритм опирается на разработанный и теоретически обоснованный в работе алгоритм вычисления вероятности, с которой модель источника ошибок смогла бы сгенерировать рассматриваемую последовательность ошибок, и представляет собой модификацию для случаев скрытой полумарковской модели Фергюсона и скрытой полумарковской QP -модели алгоритма прямого хода, описанного в работах Шенг Женг Ю.
- *в области комплексов программ:*
 - разработан и протестирован новый программный комплекс, направленный на исследование моделей источников ошибок и позволяющий решать такие задачи как генерация последовательности ошибок моделью с заданными параметрами, вычисление вероятности генерации некоторой заданной последовательности исследуемой моделью, подбор по имеющейся последовательности ошибок наиболее адекватной модели из некоторого набора.

Основные положения, выносимые на защиту.

- Скрытая полумарковская QP -модель источника ошибок в q -ичном цифровом канале передачи данных, которая обобщает ряд классических моделей

источников ошибок и которую в силу общности и особенностей структуры целесообразно впоследствии включить в ИС ОПСАПК.

- Полиномиальное представление скрытой полумарковской модели фергюссоновского типа, ускоряющее процесс генерации последовательностей ошибок, а также алгоритм его построения.
- Алгоритм численного решения задачи оценивания для общей скрытой полумарковской модели, его полное теоретическое обоснование и адаптация на случай скрытой полумарковской QP -модели и скрытой полумарковской модели фергюссоновского типа.
- Схема подбора модели источника ошибок из базы скрытых полумарковских моделей по реальной канальной последовательности и модификация информационной системы оценки применимости схем помехоустойчивого кодирования (ИС ОПСАПК) на ее основе.
- Программный комплекс, позволяющий генерировать потоки ошибок и подбирать наиболее адекватную модель источника ошибок по канальной последовательности ошибок.

Теоретическая ценность диссертационной работы состоит в разработке новых математических моделей источника ошибок в q -ичных цифровых каналах передачи данных, постановке и решении задачи адекватного представления помеховой обстановки в канале передачи данных математической моделью источника ошибок, построении и теоретическом обосновании алгоритмов подбора математической модели источника ошибок по зарегистрированной в канале последовательности ошибок, разработке алгоритма построения полиномиального представления скрытой полумарковской модели фергюссоновского типа.

Практическая ценность диссертационной работы заключается в построении модификации информационной системы оценки применимости схем помехоустойчивого кодирования, позволяющей решать нерешенную ранее задачу подбора адекватной математической модели источника для описания помеховой обстановки в исследуемом канале; разработке программного комплекса, позволяющего генерировать потоки ошибок с помощью новых моделей источников ошибок, а также реализующего несколько вариантов алгоритма подбора наиболее соответствующей модели источника ошибок по последовательности ошибок.

Реализация результатов работы.

Апробация диссертационной работы. Основные результаты работы докладывались на XI Всероссийском симпозиуме по прикладной и промышленной математике (весенняя сессия, г. Кисловодск, 1 - 8 мая 2010 г.), на международной научно-практической конференции "Информационная безопасность 2010". (г. Таганрог, 22-25 июня 2010г.), XVIII Всероссийской Школе-Коллоквиуме по стохастическим методам (весенняя сессия, Казань, 1 - 8 мая 2011 г.), на международной научно-практической конференции "Современные проблемы и достижения в отраслях радиотехники, телекоммуникаций и информационных технологий". (Украина, г. Запорожье, 19 - 21 сентября 2012 г.), на V международной

конференции "Современные методы и проблемы теории операторов и гармонического анализа и их приложения" (г. Ростов-на-Дону, 26 апреля-1 мая 2015 г.), на 4-й Российско-Китайской конференции Numerical Algebra with Applications (г. Ростов-на-Дону, 26-29 июня 2015 г.), на Международной конференции молодых ученых стран БРИКС (г. Ростов-на-Дону, 24-26 сентября 2015 г.), на VI Международной конференции "Современные методы и проблемы теории операторов и гармонического анализа и их приложения" (г. Ростов-на-Дону, 24 апреля-29 апреля 2016 г.).

Публикации. Основные результаты по теме диссертации изложены в 16 печатных изданиях [1]–[16], четыре из которых изданы в журналах, рекомендованных ВАК для публикации научных работ, отражающих основное содержание диссертации [7], [8], [10], [16], четыре — в журналах и сборниках научных работ, девять — в тезисах международных и всероссийских научных конференций. В работах [3], [5], [4], [7], [8] научному руководителю В.М. Деундяку принадлежат обсуждения формулировки результатов и методов их получения, в работах [1], [2] автору диссертации принадлежит разработка математической и компьютерной модели источника ошибок; в работах [11], [15], [16] автору диссертации принадлежит разработка метода подбора адекватной модели потока ошибок к реальному каналу связи и модификация ИС ОПСАПК на основе автоматизированного выбора модели источника ошибок.

Структура работы и объем диссертации. Работа состоит из введения, четырех глав, заключения и библиографического списка.

Во введении дается общая характеристика работы и приводятся основные результаты диссертационной работы.

Первая глава посвящена обзору методов и проблем математического моделирования цифровых помехоустойчивых каналов связи. Логически она подразделяется на два раздела. В первом разделе приведены сведения о цифровых помехоустойчивых каналах связи, имитационном моделировании таких каналов, моделях источников ошибок и задачах, возникающих при имитационном моделировании. Во втором разделе изложены основы математического аппарата теории скрытых марковских и полумарковских моделей, необходимые для дальнейшего изложения. В итоге формулируется как актуальная задача подбора математической модели помех для описания реальной помеховой обстановки.

...В работе рассматриваются q -ичные идеально синхронизированные, симметричные, нестационарные каналы с памятью.

...возникают помехи, используются кодеки,

При использовании каналов связи необходимо учитывать возможность возникновения в них ошибок. В зависимости от характеристик в реальном канале передачи данных могут возникать как одиночные ошибки, так и пакеты ошибок различной структуры и длительности. Для борьбы с ошибками применяются помехоустойчивые коды, принципом работы которых является внесение кодером избыточности в информационную последовательность и последующее использование ее при декодировании для обнаружения и исправления ошибок. Следует отметить, что различные коды и алгоритмы их декодирования неодинаково хорошо справляются с различными типами ошибок. Поэтому нужно уметь подбирать для каждого конкретного канала наиболее эффективные методы и средства обна-

ружения и исправления ошибок.

Решение этой задачи в аналитическом виде является сложным, поэтому эффективным представляется подбор помехоустойчивого кодека на основе проведения экспериментов с использованием имитационных моделей каналов с помехами и анализа их результатов. Эта задача может быть решена в рамках информационной системы оценки применимости схем алгебраического помехоустойчивого кодирования (ИС ОПСАПК).

Для функционирования информационной системы типа ИС ОПСАПК, описанной в предыдущем разделе, необходимо научиться решать задачу адекватного представления зарегистрированной в реальном канале последовательности ошибок математической моделью источника ошибок, чтобы впоследствии использовать эту модель в качестве генератора потоков ошибок при проведении имитационных экспериментов с целью подбора кодека. В силу того, что ИС ОПСАПК хранит результаты ранее проведенных экспериментов, интерес представляет не столько построение новой, адекватной каналу в смысле некоторого критерия модели источника ошибок, сколько подбор наилучшей модели из уже имеющихся в библиотеке моделей источников ошибок.

Далее будем говорить о задаче подбора модели источника ошибок как об *обратной задаче*, понимая под *прямой* задачу генерации моделью последовательности ошибок.

Представляет интерес построение моделей источника ошибок на основе скрытых полумарковских моделей, поскольку модели позволяют моделировать системы с несколькими состояниями и способны порождать последовательности наблюдений различной структуры. Кроме того, перспективным представляется использование методов статистического анализа, разработанных для скрытых полумарковских моделей, для подбора модели источника ошибок к каналу связи.

Вторая глава посвящена построению моделей источников ошибок в q -ичных цифровых каналах связи на основе скрытых марковских и полумарковских моделей. В частности, в главе строятся скрытая марковская модель источника ошибок, скрытая полумарковская QR -модель источника ошибок и скрытая полумарковская модель источника ошибок Фергюсоновского типа. Для модели Фергюсоновского типа строится полиномиальное представление. В заключительной части главы описывается программный комплекс HsmmErrorSourceGeneration, предназначенный для генерации последовательностей ошибок на основе построенных ранее в главе моделей источников ошибок.

Под **скрытой полумарковской QR-моделью** (СПМQR-моделью) источника ошибок будем понимать набор

$$\lambda = \{\mathcal{S}; \mathcal{D}; A; \Pi; p(d); \rho; \mu; M; \mathbb{F}_q; B\},$$

параметры которого определяются следующим образом.

1) $\mathcal{S} = \{1, \dots, N\}$ — алфавит состояний дискретной полумарковской цепи, $\mathcal{D} = \{1, \dots, D\}$ — алфавит возможных длительностей состояний.

2) $A = \{a_{ii'}\}_{i, i' \in \mathcal{S}}$ — матрица переходных вероятностей, матричный элемент $a_{ii'}$ которой представляют собой вероятность перехода из состояния i в состояние i' , то есть

$$\forall i, i' \in \mathcal{S} : a_{ii'} \triangleq P_\nu[S_t = i' | S_{t-1} = i],$$

где t — момент времени, когда произошёл переход из обобщённого состояния (i, d) в (i', d') . Для матрицы A должно также выполняться условие стохастичности, то есть:

$$\forall i, i' \in \mathcal{S} : 0 \leq a_{ii'} \leq 1$$

и

$$\forall i \in \mathcal{S} : \sum_{i' \in \mathcal{S}} a_{ii'} = 1.$$

3) $\Pi = \{\pi_i\}_{i \in \mathcal{S}}$ представляет собой исходное распределение вероятностей состояний. Элемент π_i — вероятность того, что система прибывала в состоянии i до начала наблюдений. Таким образом, можно записать:

$$\forall i \in \mathcal{S} : \pi_i \triangleq P[S_0 = i].$$

4) $p(d) = \{p_i(d)\}_{i \in \mathcal{S}}$ — набор распределений вероятностей длительностей в различных состояниях, где $p_i(d)$ соответствует вероятности наблюдения состояния i на протяжении в точности d моментов времени. Сказанное можно записать в виде:

$$\forall (i, d) \in \mathcal{S} \times \mathcal{D} : p_i(d) \triangleq P[S_{t+1:t+d} = i | S_{t+1} = i].$$

5) $\rho = \{\rho_i(l)\}_{i \in \mathcal{S}, l \in L_i}$ — вектор, составленный из эталонных плотностей, определяемых на эталонных отрезках $L = \{L_i\}_{i \in \mathcal{S}}$ длин $\Delta = \{\delta_i\}_{i \in \mathcal{S}}$, где

$$\forall i \in \mathcal{S}, l \in L_i : 0 \leq \rho_i(l) \leq 1$$

и

$$\forall i \in \mathcal{S} : \sum_{l \in L_i} \rho_i(l) = 1.$$

$\mu = \{\mu_i\}_{i \in \mathcal{S}}$ — вектор средних вероятностей появления ошибки в различных состояниях, а M — средняя вероятность появления ошибки в канале. Для эталонных плотностей ρ_i должно также выполняться условие адаптированности к каналу:

$$\forall i \in \mathcal{S}, \forall l \in L_i : \rho_i(l) \leq \frac{1}{\delta_i \mu_i} \leq 1.$$

6) Поле Галуа F_q — алфавит наблюдаемых символов. \mathbb{F}_q^* — мультипликативная группа поля Галуа \mathbb{F}_q . $B = \{b_i(k)\}_{i \in \mathcal{S}, k \in \mathbb{F}_q^*}$ — набор распределений вероятностей значений ошибки в состояниях, элемент которого $b_i(k)$ представляет собой вероятность наблюдать символ k в состоянии i , т.е. можно записать в виде:

$$\forall i \in \mathcal{S}, \forall k \in \mathbb{F}_q^* :$$

$$b_i(k) \triangleq P[k | S_t = i].$$

Посредством описанной модели можно производить генерацию потоков ошибок. В нулевой момент времени (то есть до начала генерации последовательности ошибок) с помощью вектора Π определяется состояние системы до наступления первого момента времени. Далее для каждого момента времени выполняется следующая последовательность действий:

1) На основе матрицы переходных вероятностей выбирается текущее состояние канала i .

2) С использованием распределения вероятностей длительностей $p_i(d)$ в выбранном состоянии определяется длина d временного отрезка, на протяжении которого модель будет находиться в текущем состоянии (то есть длительность текущего состояния).

3) Для каждой точки отрезка выбранной длины вычисляется вероятность наличия в ней ошибки:

$$\forall \theta \in [1, d] : e_{i,\theta}^d = \varphi_i^d(\theta) \mu_i d, \quad (1)$$

где

$$\varphi_i^d(\theta) = \int_{\theta}^{\theta+1} \varphi_{i,d}^{pc}(x) dx \quad (2)$$

и

$$\forall l \in L_i : \quad \varphi_{i,d}^{pc} |_{(lq, (l+1)q]} = \frac{\delta_i}{d} \cdot \rho(i),$$

где $q = \frac{d}{\delta_i}$. В соответствии с этой вероятностью определяется наличие ошибки на данной позиции.

4) В случае, если ошибка есть, в соответствии с матрицей B генерируется ее значение. В противном случае, на этой позиции будет стоять нулевой элемент.

Скрытой полумарковской моделью фергюсоновского типа будем называть систему:

$$\Delta = \{S, A, \pi, V, B, D, F\}, \quad (3)$$

где 1) $S = \{S_1, S_2, \dots, S_N\}$ — алфавит состояний модели;

2) $A = \{a_{ij}\}_{i,j=1}^N$ — матрица переходов;

3) $\pi = \{\pi_i\}_{i=1}^N$ — вектор начального распределения вероятностей состояний;

4) $V = \{v_1, v_2, \dots, v_M\}$ — алфавит наблюдаемых символов;

5) $B = \{b_{ij}\}_{i,j=1}^{N,M}$ — матрица условных вероятностей наблюдения символа v_j в состоянии S_i ;

6) D — алфавит возможных длительностей мощности L ;

7) $F = \{f_{ij}\}_{i,d=1}^{N,L}$ — матрица размера $N \times L$, элемент f_{ij} которой представляет собой вероятность наблюдать длительность d_j при условии, что Δ находится в состоянии S_i .

Процесс генерирования последовательности наблюдений посредством модели типа Фергюсона выглядит следующим образом.

Этап 1. В соответствии либо с вектором π (в момент времени $t = 1$), либо с матрицей A выбирается текущее состояние.

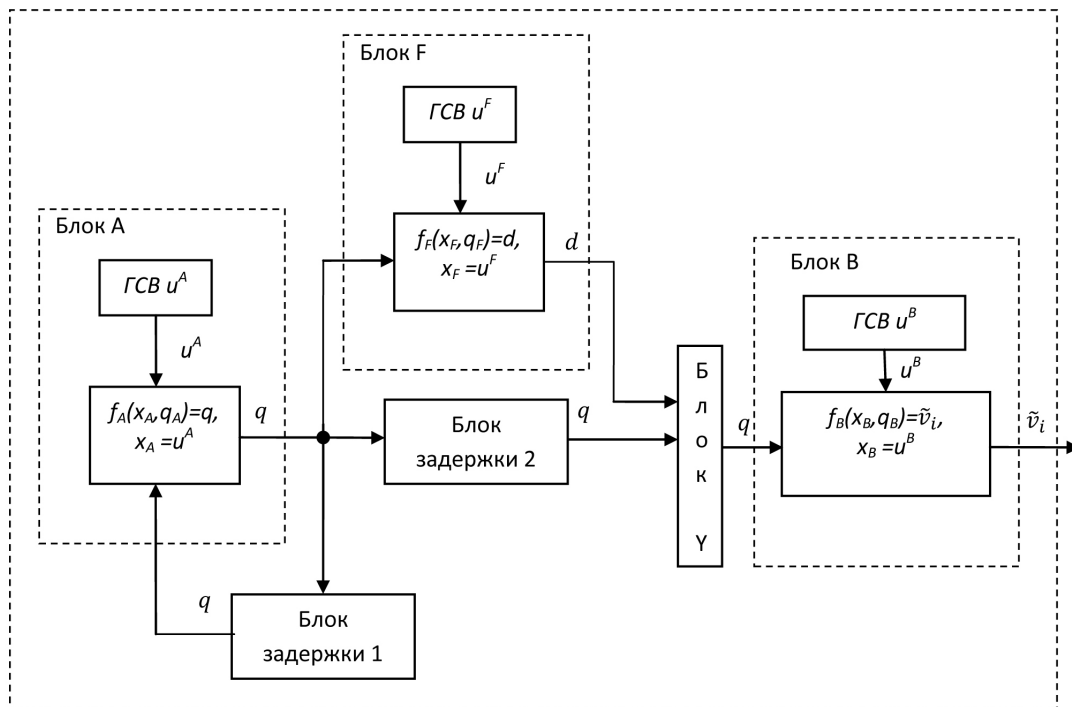
Этап 2. В выбранном состоянии S_i с использованием $p_i(d)$ определяется длина временного промежутка, в течение которого система будет находиться в i -ом состоянии.

Этап 3. В каждой точке этого отрезка посредством $b_i(k)$ генерируется наблюдение.

Полиномиальной моделью скрытой полумарковской модели типа Фергюсона назовем систему

$$\text{ПМ} = (u^A, u^B, u^F, f_B(x_B, f_A(x_A, q_A)), f_F(x_F, f_A(x_A, q_A)), \pi),$$

в которой u^A, u^B и u^F — случайные величины, определяемые по разложениям матриц A , B и F на линейные комбинации простых матриц, значения которых закодированы элементами поля F_{2^n} . Структурная схема автомата, реализующего полиномиальное представление скрытой полумарковской модели фергюсоновского типа Δ , представлена на рисунке .



Рис

1. Структурная схема автомата, реализующего полиномиальное представление скрытой полумарковской модели типа Фергюсона.

Для генерации потоков ошибок скрытой полумарковской QP -моделью и скрытой полумарковской моделью Фергюсоновского типа разработан специальный комплекс программ **HsmmErrorSourcesGeneration**. Основной задачей этого программного комплекса является предоставление исследователю возможности осуществлять генерацию последовательностей ошибок в соответствии с заданными в виде текстового файла параметрами модели источника ошибок и алгоритмом генерации для модели заданного типа. Полученные на выходе текстовые файлы, содержащие сгенерированные последовательности, могут использоваться исследователем для самостоятельного анализа.

Программный комплекс **HsmmErrorSourcesGeneration** реализован на языке программирования высокого уровня $C\sharp$. Необходимым условием для работы программного продукта является наличие .NET Framework 4.0. В основе архитектуры комплекса лежат принципы объектно-ориентированного дизайна.

В третьей главе предлагаются алгоритмы решения задачи оценивания в случае общей скрытой полумарковской модели, скрытой полумарковской модели Фергюсона и скрытой полумарковской QP -модели, а также задачи подбора модели источника ошибок из набора моделей, позволяющей наилучшим образом генерировать заданную последовательность ошибок.

Эта задача не имеет решения в общем виде, но может быть решена для скрытых марковских и полумарковских моделей. В этом случае ее решение основывается на решении *задачи оценивания*, т.е. на вычислении вероятности генерации последовательности ошибок заданной скрытой полумарковской моделью. В работе Рабинера предлагается алгоритм решения задачи оценивания для скрытых марковских моделей, а Ю предлагает ее решение для общих скрытых полумарковских моделей при дополнительных предположениях относительно начала и конца скрытой последовательности состояний, а именно, что первое наблюдаемое состояние началось в или до первого наблюдаемого момента времени, а последнее состояние закончилось строго в момент времени T .

В главе рассматриваются и теоретически обосновываются два варианта решения задачи оценивания для случая общей скрытой полумарковской модели при более общих, чем у Ю, предположениях относительно начала и конца наблюдаемой части последовательности состояний.

Теорема 1. Вероятность генерации последовательности $O_{1:T}$ общей скрытой полумарковской моделью λ при общем предположении может быть вычислена по следующей формуле:

$$\tilde{P}[O_{1:T}] = \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:T-d_1}] \bar{\alpha}_{T-d_1+d}(j, d) b_{j,d}(O_{T-d_1+1}^{T-d_1+d}),$$

где

$$\begin{aligned} \tilde{P}[O_{1:t}] &= \begin{cases} 1, & t \leq 0, \\ \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:t-d_1}] \bar{\alpha}_{t-d_1+d}(j, d) b_{j,d}(O_{t-d_1+1}^{t-d_1+d}), & t \in [1, T], \end{cases} \\ \bar{\alpha}_t(i, d) &= \begin{cases} \pi_{i,d}, & t \leq 0, \\ \sum_{i' \in \mathcal{S}} \sum_{d' \in \mathcal{D}} \bar{\alpha}_{t-d}(i', d') \bar{b}_{i',d'}(O_{t-d-d'+1}^{t-d-d'}) a_{(i',d')(i,d)}, & t > 0, \end{cases} \\ \bar{b}_{i',d'}(O_{t-d-d'+1}^{t-d-d'}) &= \frac{b_{i',d'}(O_{t-d-d'+1}^{t-d-d'}) \tilde{P}[O_{1:t-d-d'}]}{\tilde{P}[O_{1:t-d}]}. \end{aligned}$$

Теорема 2. Будем предполагать, что для последовательности $O_{1:T}$ вероятность последнего наблюдаемого состояния удовлетворяет предельному распределению вероятностей. Тогда вероятность генерации последовательности $O_{1:T}$ общей скрытой полумарковской моделью λ при общем предположении может быть вычислена по следующей формуле:

$$\tilde{P}[O_{1:T}] = \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \frac{\pi_{j,d}}{d} \sum_{d_1=1}^d P[O_{1:T-d_1}] b_{j,d}(O_{T-d_1+1}^T), \quad (4)$$

где

$$P[O_{1:t}] = \begin{cases} 1, & t \leq 0, \\ \sum_{i \in \mathcal{S}} \sum_{d \in \mathcal{D}} P[O_{1:t-d}] \bar{\alpha}_t(i, d) b_{i,d}(O_{t-d+1:t}), & 0 < t \leq T - d_1, \end{cases} \quad (5)$$

$$\bar{\alpha}_t(i, d) = \begin{cases} \pi_{i,d}, & t \leq 0, \\ \sum_{i' \in \mathcal{S}} \sum_{d' \in \mathcal{D}} \bar{\alpha}_{t-d}(i', d') \bar{b}_{i',d'}(O_{t-d-d'+1}^{t-d-d'}) a_{(i',d')(i,d)}, & t > 0, \end{cases} \quad (6)$$

$$\bar{b}_{i',d'}(O_{t-d-d'+1}^{t-d}) = \frac{b_{i',d'}(O_{t-d-d'+1}^{t-d})P[O_{1:t-d-d'}]}{P[O_{1:t-d}]}$$

На основе теоремы 1 в главе строятся решения задачи оценивания для скрытой полумарковской модели Фергюсона и скрытой полумарковской QP -модели.

Теорема 3. Вероятность генерации последовательности $O_{1:T}$ скрытой полумарковской моделью Фергюсона λ при общем предположении может быть вычислена по следующей формуле:

$$\begin{aligned} \tilde{P}[O_{1:T}] &= \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:T-d_1}] \bar{\alpha}_{T-d_1+d}(j, d) \prod_{\tau=T-d_1+1}^{T-d_1+d} b_j(O_\tau), \\ \tilde{P}[O_{1:t}] &= \begin{cases} 1, & t \leq 0, \\ \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:t-d_1}] \bar{\alpha}_{t-d_1+d}(j, d) \prod_{\tau=t-d_1+1}^{t-d_1+d} b_j(O_\tau), & t \in [1, T], \end{cases} \\ \bar{\alpha}_t(i, d) &= \begin{cases} \pi_i \rho_i(d), & t \leq 0, \\ \sum_{i' \in \mathcal{S}} \sum_{d' \in \mathcal{D}} \bar{\alpha}_{t-d}(i', d') \prod_{\tau=t-d-d'+1}^{t-d} a_{i'i} p_i(d) b_{i'}(O_\tau) \frac{\tilde{P}[O_{1:t-d-d'}]}{\tilde{P}[O_{1:t-d}]}, & t \in [1, T]. \end{cases} \end{aligned}$$

Теорема 4. Вероятность генерации последовательности $O_{1:T}$ скрытой полумарковской QP -моделью λ при общем предположении может быть вычислена по следующей формуле:

$$\tilde{P}[O_{1:T}] = \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:T-d_1}] \bar{\alpha}_{T-d_1+d}(j, d) \prod_{\tau=T-d_1+1}^{T-d_1+d} b_{i,d}^\tau(O_\tau),$$

где

$$\begin{aligned} \tilde{P}[O_{1:t}] &= \begin{cases} 1, & t \leq 0, \\ \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:t-d_1}] \bar{\alpha}_{t-d_1+d}(j, d) \prod_{\tau=t-d_1+1}^{t-d_1+d} b_{i,d}^\tau(O_\tau), & t \in [1, T], \end{cases} \\ \bar{\alpha}_t(i, d) &= \begin{cases} \pi_i p_i(d), & t \leq 0, \\ \sum_{i' \in \mathcal{S}} \sum_{d' \in \mathcal{D}} \bar{\alpha}_{t-d}(i', d') \prod_{\tau=t-d-d'+1}^{t-d} b_{i',d'}^\tau(O_\tau) \frac{P[O_{1:t-d-d'}]}{P[O_{1:t-d}]} a_{i'i} p_i(d), & t > 0, \end{cases} \\ b_{i,d}^\tau(O_\tau) &= I_{\mathbb{F}_q^*}^\tau(O_\tau) \varphi_i^d(\tau) \mu_i d b_i(O_\tau) + (1 - I_{\mathbb{F}_q^*}^\tau(O_\tau)) (1 - \varphi_i^d(\tau) \mu_i d). \end{aligned}$$

Далее в главе приводятся формальные описания алгоритмов решения задачи оценивания и задачи подбора по последовательности ошибок математической модели источника ошибок из класса скрытых полумарковских моделей.

Приведем алгоритм решения задачи оценивания на примере скрытой полумарковской QP модели.

Data: $O = O_1O_2...O_T$ – последовательность ошибок, λ – скрытая полумарковская модель источников ошибок.

Result: вероятность генерации последовательности O моделью λ – $P(O|\lambda)$.

Вычислить $P(O|\lambda) = \tilde{P}[O_{1:T}]$ по следующим формулам:

$$\tilde{P}[O_{1:T}] = \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d P[O_{1:T-d_1}] \bar{\alpha}_{T-d_1+d}(j, d) b_{j,d}(O_{T-d_1+1}^{T-d_1+d}),$$

где

$$\tilde{P}[O_{1:t}] = \begin{cases} 1, & t \leq 0, \\ \sum_{j \in \mathcal{S}} \sum_{d \in \mathcal{D}} \sum_{d_1=1}^d \tilde{P}[O_{1:t-d_1}] \bar{\alpha}_{t-d_1+d}(j, d) \prod_{\tau=t-d_1+1}^{t-d_1+d} b_{j,d}^{\tau}(O_{\tau}), & t \in [1, T], \end{cases}$$

$$\bar{\alpha}_t(i, d) = \begin{cases} \pi_i p_i(d), & t \leq 0, \\ \sum_{i' \in \mathcal{S}} \sum_{d' \in \mathcal{D}} \bar{\alpha}_{t-d}(i', d') \prod_{\tau=t-d-d'+1}^{t-d} b_{i',d'}^{\tau}(O_{\tau}) \frac{P[O_{1:t-d-d'}]}{P[O_{1:t-d}]} a_{i'i} p_i(d), & t > 0, \end{cases}$$

$$b_{i,d}^{\tau}(O_{\tau}) = I_{\mathbb{F}_q^*}(O_{\tau}) \varphi_i^d(\tau) \mu_i d b_i(O_{\tau}) + (1 - I_{\mathbb{F}_q^*}(O_{\tau})) (1 - \varphi_i^d(\tau) \mu_i d).$$

return $P(O|\lambda) = \tilde{P}[O_{1:T}]$

Algorithm 1: EvaluationProblemSolver

Эксперименты показывают, что использование апостериорных вероятностей полностью не решает проблему антипереполнения. Поэтому далее в главе предлагается усредняющий алгоритм, в основе которого лежит деление исходной последовательности на части и усреднение вероятностей наблюдения этих частей.

Data:

- 1) последовательность ошибок $O = O_1O_2...O_T$;
- 2) скрытая полумарковская модель λ ;
- 3) длина сегмента разбиения $l \leq T$.

Result: усредненная вероятность генерации последовательности O моделью λ .

Разбить O на $n = \lfloor T/l \rfloor$ частичных последовательностей $\{O_i\}_{i=1}^n$, где под $\lfloor T/l \rfloor$ будем понимать ближайшее целое число, не превосходящее T/l .

for *всех* i *из* $(0, n]$ **do**

вычислить вероятность $P(O_i|\lambda) = \tilde{P}[O_i|\lambda]$ в соответствии с алгоритмом InverseProblemSolver

end

Вычислить усредненную вероятность $P_{avg}(O|\lambda) = \frac{\sum_{i \in (0, n]} P(O_i|\lambda)}{n}$, **return** $P_{avg}(O|\lambda)$.

Algorithm 2: EvaluationProblemSolverExtended

Алгоритм решения задачи подбора выглядит следующим образом:

В заключительной части главы представлено описание и особенности реализации программного комплекса **HsmmErrorSourcesAnalysis**, позволяющего подбирать модель источника из класса скрытых полумарковских QP -моделей и скры-

Data:

- 1) последовательность ошибок $O = O_1 O_2 \dots O_T$;
- 2) база Λ , составленная из скрытых полумарковских моделей источников ошибок с различными параметрами.

Result: модель λ_{max} , позволяющая наилучшим образом имитировать ошибку в канале.

for *все* $\lambda \in \Lambda$ **do**

вычислить вероятность $P_{avg}(O|\lambda)$ в соответствии с алгоритмом
EvaluationProblemSolverExtended

end

Найти $P_{max} = \max_{\lambda \in \Lambda} P(O|\lambda)$.

Выбираем модель, такую что

$$\lambda_{max} = \arg \max_{\lambda_m \in \Lambda} P(O|\lambda) = \arg P_{max},$$

где функция \arg возвращает модель λ_{max} , соответствующую максимальной вероятности P_{max} . **return** λ_{max} .

Algorithm 3: InverseProblemSolver

тых полумарковских моделей Фергюсоновского типа, наиболее адекватно подходящую для генерации заданной на вход последовательности.

Данный программный комплекс предоставляет исследователю возможность рассчитывать вероятность генерации заданной пользователем последовательности выбранной скрытой полумарковской моделью одного из перечисленных выше типов, а также выбирать наиболее подходящую в смысле выбранного критерия модель по заданной последовательности.

Программный комплекс **HsmmErrorSourcesAnalysis** реализован на языке программирования высокого уровня C#. Необходимым условием для работы программного продукта является наличие .NET Framework 4.0. В основе архитектуры комплекса лежат принципы объектно-ориентированного дизайна.

Четвертая глава посвящена применению разработанных методов решения обратных задач на практике. В первой части главы строится модификация ИС ОПСАПК на основе разработанного алгоритма подбора модели источника ошибок по канальной последовательности ошибок. Во второй части главы подробно рассматривается разработанный программный комплекс для моделирования и исследования потоков ошибок.

Рассмотрим информационную систему оценки применимости схем алгебраического помехоустойчивого кодирования (ИС ОПСАПК), база источников ошибок (БМИО) в которой содержит только модели из класса скрытых полумарковских моделей. Расширим функции этой информационной системы возможностью анализировать приходящую канальную последовательность ошибок и находить для нее такую модель источника ошибок из БМИО, которая способна генерировать последовательности ошибок, близкие к полученной последовательности.

Для решения этой задачи расширим блок управления имитационной моделью (БУИМ) из ИС ОПСАПК модулем подбора математической модели источника ошибок (МПМИО), построенного на основе разработанного в главе 3 алгоритма.

Задача МПМИО состоит в выборе из набора моделей источников ошибок модели, наилучшим образом соответствующей реальной канальной последовательности. Расширенный блок обозначим БУИМ(А)

Схема потоков данных между элементами модифицированной информационной системы представлена на рисунке 1.

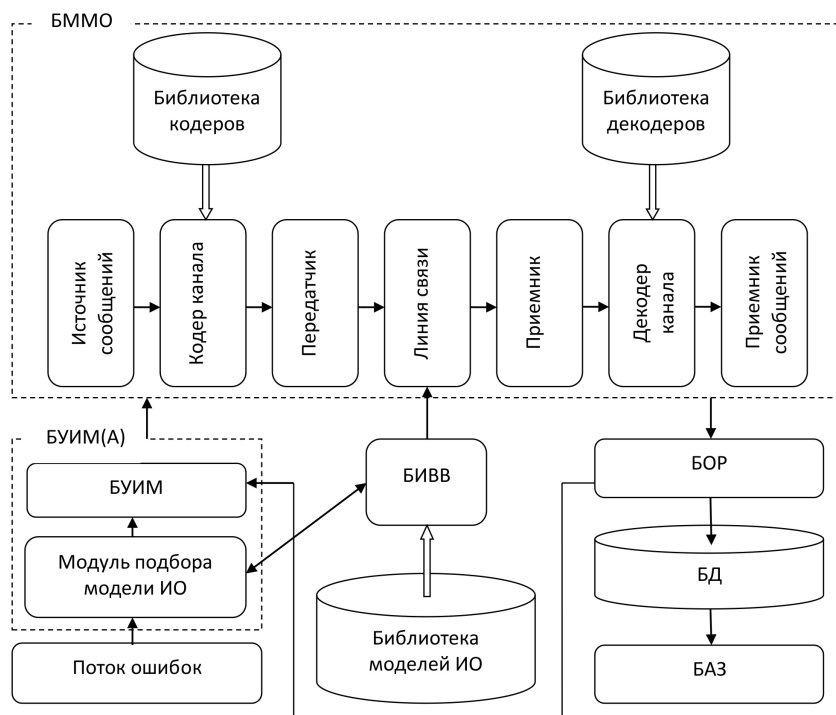


Рис.1 Модифицированная схема ИС ОПСАПК.

На вход БУИМ(А) подается последовательность наблюдений $O_{1:T}$. БУИМ передает эту последовательность МПМИО, который в соответствии с некоторым критерием K определяет подмножество моделей источника ошибок $\Lambda_K = \{\lambda_i\}$ из БМИО, способных сгенерировать такую последовательность. Далее МПМИО с помощью метода подбора из предыдущего раздела выбирает из него наиболее адекватную модель λ . Выбранная модель λ возвращается в БУИМ, который запрашивает у БОР результаты R_λ интересующих пользователя экспериментов, проведенных с участием этой модели. Если искомые результаты R_λ найдены, то необходимость в проведении имитационных экспериментов отпадает. Если же эксперименты для выбранной модели еще не проводились, то модель λ подается блоком управления на вход генератора потоков ошибок, который производит генерацию потока ошибок моделью λ , и БУИМ запускает имитационные эксперименты.

Далее в главе 4 описывается разработанный программный комплекс для моделирования и анализа потоков ошибок. Комплекс предоставляет пользователю удобный интерфейс

- [1] Бибов А.Ю. , Жданова М. А. Экспериментальное исследование аналога списочного декодера Сидельникова с использованием обобщенной марковской модели источника ошибок. // Информационное противодействие угрозам терроризма – 2009. – № 13. – С. 151-153.
- [2] Бибов А.Ю. , Жданова М. А. Экспериментальное исследование аналога списочного декодера Сидельникова с использованием обобщенной марковской модели источника ошибок. // Материалы XI Международной научно-практической конференции «Информационная безопасность», часть 1, Таганрог: Изд-во ТТИ ЮФУ, 2010, с. 221-225.
- [3] Деундяк В.М., Жданова М.А. Обобщенная марковская модель источника ошибок q-ичного цифрового канала нескольких физических состояний. // Математика и ее приложения: ЖИМО. – Иваново: ИвГУ. – 2010. – Выпуск 1 (7). – С.34-40.
- [4] Деундяк В.М., Жданова М.А. О некотором обобщении марковской математической модели источника ошибок. // Обзорение прикладной и промышленной математики. – 2010, т.17, №5, с.713-714.
- [5] Деундяк В.М., Жданова М.А. О применении скрытых марковских моделей в моделировании источников ошибок. // Обзорение прикладной и промышленной математики. – 2011. –Выпуск 3. –С. 488.
- [6] Деундяк В.М., Жданова М.А. Об аппроксимации потока ошибок в канале передачи данных на основе скрытых полумарковских QR-моделей // Тези доповідей VI Міжнародної науково-практичної конференції «Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій», 19-21 вересня 2012 р., м. Запоріжжя, 2012, С.109-110.
- [7] Деундяк В.М., Жданова М.А. Полиномиальное представление скрытой полумарковской модели фергюсоновского типа. // Вестник ВГУ, Серия: Системный анализ и информационные технологии – 2013, №2, с.71-78.
- [8] Деундяк В.М., Жданова М.А. Решение задачи оценивания скрытых полумарковских QR-моделей. // Вестник ДГТУ, Т.14, № 4(79), 2014, с.5-16.
- [9] Жданова М.А. Экспериментальное исследование решения задачи оценивания скрытой полумарковской QR-модели // Материалы всероссийской научно-практической школы-конференции по информационной безопасности «Вестник по безопасности», выпуск №7.-Тольятти, ВуиТ, 2014, с. 10-11.
- [10] Деундяк В.М., Жданова М.А. О решении задачи оценивания скрытых полумарковских моделей фергюсоновского типа. // Известия вузов. Сев.-Кавк. Регион. Естественные науки. 2015. N 3, с. 19-24.
- [11] Деундяк В.М., Жданова М.А., Могилевская Н.С. Об автоматизированном выборе модели потока ошибок в информационной системе оценки применимости помехоустойчивого кодирования // Труды научной школы И.Б.Симоненко. Выпуск 2. Ростов-на-Дону: издательство ЮФУ, 2015. с. 111-120.

- [12] *Zhdanova M.A.* Inverse problems of HSMM-based mathematical modeling of jamming environment // Международная научная конференция «Современные методы и проблемы теории операторов и гармонического анализа и их приложения VI». Ростов-на-Дону, 26 апреля - 1 мая 2015г. Материалы конференции. Изд. Центр ДГТУ, Ростов-на-Дону, 2015, с.185.
- [13] *Zhdanova M.A., Deundyak V.M.* Evaluation problem for general hidden semi-Markov error source model// Numerical algebra with applications. Procceding of 4 China-Russia conference. Rostov-on-Don: SFUP, 2015. P.156-159.
- [14] *Жданова М.А.* Об экспериментальном исследовании решения задачи оценивания некоторых скрытых полумарковских моделей. // Международной конференции молодых ученых стран БРИКС «Сотрудничество стран БРИКС для устойчивого развития», Ростов-на-Дону, 24 – 26 сентября 2015г. Материалы конференции. с. XXX-XXX.
- [15] *Деундяк В.М., Жданова М.А., Могилевская Н.С.* Автоматизированный выбор модели потока ошибок в информационной системе оценки применимости помехоустойчивого кодирования // Международная научная конференция «Современные методы и проблемы теории операторов и гармонического анализа и их приложения VI». Ростов-на-Дону, 24 – 29 апреля 2016г. Материалы конференции. Изд. Центр ДГТУ, Ростов-на-Дону, 2016, с.157-158
- [16] *Деундяк В.М., Жданова М.А., Могилевская Н.С.* Решение задачи подбора модели источника ошибок в ИС ОПСАПК. // Вестник Донского государственного технического университета. 2017. №4(91), с. 107-115.