

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение высшего профессионального образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

В.Г. Букреев, С.И. Колесникова, А.Е. Янковская

**ВЫЯВЛЕНИЕ ЗАКОНОМЕРНОСТЕЙ
ВО ВРЕМЕННЫХ РЯДАХ В ЗАДАЧАХ
РАСПОЗНАВАНИЯ СОСТОЯНИЙ
ДИНАМИЧЕСКИХ ОБЪЕКТОВ**

Монография

Издательство
Томского политехнического университета
2010

УДК 519.246

ББК 22.17

Б90

Букреев В.Г.

Б90 Выявление закономерностей во временных рядах в задачах распознавания состояний динамических объектов: монография / В.Г. Букреев, С.И. Колесникова, А.Е. Янковская; Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2010. – 254 с.

ISBN 978-5-98298-694-8

В монографии представлены новые информационные модели распознавания состояний слабоформализованных динамических объектов, полученные на базе математического аппарата теории распознавания образов, теории вероятностей и математической статистики, теории информации, формализма мультимножеств и теории принятия решений. Модели синтезированы на основе выявления закономерностей в стохастических временных рядах, сопровождающих поведение управляемых динамических объектов в пространстве измеряемых переменных состояния.

Работоспособность моделей и алгоритмов распознавания состояний динамических объектов показана при решении ряда прикладных задач, в том числе, для организации адаптивного управления сложными объектами.

Предназначена для специалистов в области искусственного интеллекта и автоматизированного принятия решений, а также для широкого круга научных работников, аспирантов и студентов вузов технических специальностей.

УДК 519.246

ББК 22.17

Рецензенты

Доктор технических наук, профессор ТГУ

В.В. Поддубный

Доктор технических наук, профессор ТУСУРа

А.А. Светлаков

ISBN 978-5-98298-694-8

© ГОУ ВПО НИ ТПУ, 2010

© Букреев В.Г., Колесникова С.И.,
Янковская А.Е., 2010

© Обложка. Издательство Томского
политехнического университета, 2010

ОГЛАВЛЕНИЕ	3
ВВЕДЕНИЕ В ПРОБЛЕМУ	6
Глава 1. ОБЗОР МЕТОДОВ И ФОРМАЛИЗАЦИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ СОСТОЯНИЙ СЛОЖНОГО ДИНАМИЧЕСКОГО ОБЪЕКТА.....	11
1.1. Обзор современных методов и информационных технологий распознавания состояний сложного динамического объекта	11
1.2. Основные понятия и определения	20
1.3. Формализация задачи распознавания состояний динамического объекта	25
1.4. Пример информационной модели оценивания состояний сложного объекта на основе временных рядов	28
Глава 2. ЭТАЛОННАЯ МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ СОСТОЯНИЙ СЛОЖНОГО ДИНАМИЧЕСКОГО ОБЪЕКТА	33
2.1. Общая схема оценивания состояний ДО по временным рядам методами распознавания образов и нелинейной динамики	33
2.2. Метод обобщенных эталонов и решение задачи распознавания состояний динамического объекта	36
2.3. Реконструкция формы состояний стохастического объекта на основе модели эталонов	47
2.3.1. Метод моделирования состояний ДО на основе обобщенных эталонов и параметризации модели авторегрессии	51
2.4. Применение алгебраического подхода к анализу зашумленных временных рядов.....	59
2.4.1. Алгоритм реконструкции нестационарного временного ряда на основе обобщенных эталонов и модификации метода Бока	64
2.5. Идентификация состояний ДО на основе непараметрической нелинейной регрессионной модели временного ряда	67
2.6. Метод цифрового дифференцирования зашумленного временного ряда на основе модели состояний ДО	70
2.7. Определение типа аттрактора и времени вхождения в аттрактор	76
2.8. Численное моделирование распознавания состояний ДО и оценка его качества	77
Глава 3. МОДИФИКАЦИЯ МЕТОДА АНАЛИЗА ИЕРАРХИЙ ДЛЯ ДИНАМИЧЕСКИХ НАБОРОВ АЛЬТЕРНАТИВ	80
3.1. История вопроса. Нежелательные эффекты МАИ.....	81
3.2. Основные определения и понятия. Постановка задачи	82
3.3. Нелинейная модификация МАИ и ее свойства	87
3.4. Иллюстративный пример применения ММАИ	89
Глава 4. МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ В ДИСКРЕТНОМ ПОДХОДЕ К РАСПОЗНАВАНИЮ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА ...	93
4.1. Основные понятия и определения	93
4.2. Формализм мультимножеств	98
4.3. Алгоритм тестового распознавания состояния ДО	100
4.4. Алгоритм распознавания состояния ДО на основе формализма мультимножеств.....	105
4.5. Методы оценивания весовых коэффициентов признаков в тестовом распознавании образов.....	106
4.6. Дискретный и статистический подходы к оцениванию весовых коэффициентов	

признаков и тестов.....	107
4.7. Метод оценивания весовых коэффициентов признаков и тестов на основе формализма мультимножеств и нелинейной модификации метода анализа иерархий	112
4.8. Метод оценивания весовых коэффициентов признаков и тестов при большой размерности признакового пространства	114
4.8.1. Сравнительный пример определения ВКП по разным методам	117
4.9. Показатели эффективности методов оценивания признаков, тестов и принятие итогового решения.....	119
Глава 5. МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ В ТЕОРЕТИКО-ИНФОРМАЦИОННОМ ПОДХОДЕ К РАСПОЗНАВАНИЮ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА	124
5.1. Оптимальная оценка вероятности состояний сложного объекта на основе максимального правдоподобия	124
5.2. Итеративная процедура для выбора системы признаков на основе максимального правдоподобия	127
5.3. Теоретико-информационный подход к выявлению закономерностей в задаче распознавания состояний динамического объекта	128
5.4. Оценка информативности квантования временного ряда в задаче распознавания состояний динамического объекта	130
5.4.1. Практическая реализация адаптивного выбора эффективной системы признаков.....	131
5.5. Численное сравнение эффективности равномерного и адаптивного квантования при распознавании состояний ЭМС.....	138
Глава 6. ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ РАСПОЗНАВАНИЯ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА	142
6.1. Постановка задачи оценки качества процедур распознавания состояний ДО.....	143
6.2. Вероятностная модель пересечения случайным процессом определенного уровня.....	144
6.3. Двусторонние оценки для вероятности достижения гарантированного уровня качества процедуры распознавания	148
Глава 7. АДАПТИВНОЕ УПРАВЛЕНИЕ СЛАБОФОРМАЛИЗОВАННЫМ ДИНАМИЧЕСКИМ ОБЪЕКТОМ НА ОСНОВЕ МОДЕЛИ СОСТОЯНИЯ.....	151
7.1. Постановка задачи синтеза управления по методу АКАР	152
7.2. Постановка задачи синтеза управления ДО в пространстве состояний	154
7.3. Регулятор для управления слабоформализованным ДО в пространстве состояний	155
7.4. Регулятор для управления слабоформализованным ДО в пространстве наблюдений	167
7.5. Аналитическое конструирование нелинейных систем с наблюдателем состояния на основе развития АКАР	168
7.6. Иллюстративный пример построения наблюдателя для неизменяемых координат и возмущений	173
Глава 8. МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ДЛЯ РАСПОЗНАВАНИЯ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА	180
8.1. Анализ современного состояния интеллектуальных систем для распознавания состояний динамических объектов	180
8.2. Концептуальная схема и структура инструментального средства ИИС ИМСЛОГ.....	182

8.2.1. Программная реализация методов определения весовых коэффициентов в ИИС ИМСЛОГ.....	187
8.3. Методология распознавания состояний динамических объектов в интеллектуальной системе IReDSS	188
Глава 9. РЕШЕНИЕ ПРИКЛАДНЫХ ЗАДАЧ	192
9.1. Прикладная задача построения модели наблюдателя при адаптивном управлении электро-механическим объектом	192
9.1.1. Задача управления и синтеза наблюдателя состояния нелинейного электромеханического объекта на основе модели состояния	193
9.1.2. Синтез непрерывного векторного закона управления асинхронным двигателем	195
9.1.3. Конструирование наблюдателя частоты вращения ротора асинхронным двигателем ..	198
9.2. Прикладная задача экологического диагностирования и прогнозирования на основе временных рядов и обобщенных эталонов.....	201
9.3. Прикладная задача моделирования состояния здоровья пациента с бронхиальной астмой.....	207
9.4. Прикладная задача предсказания неизвестных значений непрерывных атрибутов в базах данных.....	215
ЗАКЛЮЧЕНИЕ	219
ЛИТЕРАТУРА.....	221
ПРИЛОЖЕНИЕ к главе 2	234
Доказательство утверждения 2.1.....	234
Доказательство утверждения 2.2.....	234
Доказательство утверждения 2.3.....	235
ПРИЛОЖЕНИЕ к главе 3	235
Доказательство утверждения 3.1.....	235
Доказательство теоремы 3.1.....	236
Доказательство теоремы 3.2.....	237
Доказательство теоремы 3.3.....	238
Доказательство теоремы 3.4.....	238
ПРИЛОЖЕНИЕ к главе 4	239
Доказательство теоремы 4.1.....	239
Доказательство теоремы 4.2.....	240
Доказательство теоремы 4.3.....	241
Иллюстративный пример применения формализма мультимножеств	242
Алгоритм П4.1. Поиск всех кратчайших столбцовых покрытий двоичной матрицы.....	243
ПРИЛОЖЕНИЕ к главе 5	245
Доказательство теоремы 5.2.....	245
Доказательство леммы 5.1.....	246
Доказательство леммы 5.2.....	247
Доказательство теоремы 5.3.....	248
ПРИЛОЖЕНИЕ к главе 6	249
Доказательство леммы 6.3.....	249
Доказательство леммы 6.4.....	249
Доказательство теоремы 6.1.....	250
Обозначения и сокращения	253

ВВЕДЕНИЕ В ПРОБЛЕМУ

Нельзя построить содержательную общую теорию нелинейных систем
/Джон фон Нейман/

В последние годы все более актуальным становится вопрос об оценке текущего состояния промышленных, экологических, геофизических и других объектов. Так, методы и средства оценки технического состояния машин и оборудования включают контроль (получение информации о величинах параметров и зонах их допустимого отклонения), мониторинг (информация о тенденциях изменения параметров во времени), диагностирование (идентификацию места, вида и величины дефекта). Наиболее сложна задача прогноза развития дефекта, а не изменений контролируемых параметров, решение которой позволяет определить остаточный ресурс или прогнозируемый интервал безаварийной работы. В настоящее время под термином мониторинг часто понимается решение всего комплекса процедур оценки состояния. При принятии решений, как правило, анализируемый динамический объект (ДО) является слабо структурированным, математической формализации поддаются отдельные фрагменты общего описания. Объекты, не имеющие полного аналитического описания и образующие во времени многомерный нестационарный случайный процесс (при дискретизации – слабопредсказуемый временной ряд), называются сложными ДО. На практике большинство процессов являются нестационарными, а все реальные динамические объекты - нелинейными [1, 3, 26, 32, 35, 37, 51, 117, 129, 135, 159].

В настоящее время понятия динамического объекта и динамической системы раскрываются весьма широко и по-разному трактуются разными авторами. Будем придерживаться следующего [37]: ДО – это объекты (системы) любой природы: физической, химической, биологической, экономической и др., причем не только детерминированные системы, но и стохастические, состояние которой изменяется (дискретно или непрерывно) во времени. По тексту работы различия между ДО и динамической системой проводиться не будет.

Задача поддержания работоспособного состояния ДО заключается в своевременном распознавании критических состояний ДО и устранении причин их появления (рис. 0.1) [118]. На рис. 0.2 помещен план книги, отражающий основную суть обсуждаемой в книге схемы распознавания состояний ДО, реализованной новыми и адаптированными существующими методами и алгоритмами для решения задачи распознавания состояний ДО.

Приведем несколько примеров из разных предметных областей, в том числе и рассмотренных в книге постановок прикладных задач (гл. 9), свидетельствующих об актуальности вопроса оценки состояния динамических объектов.

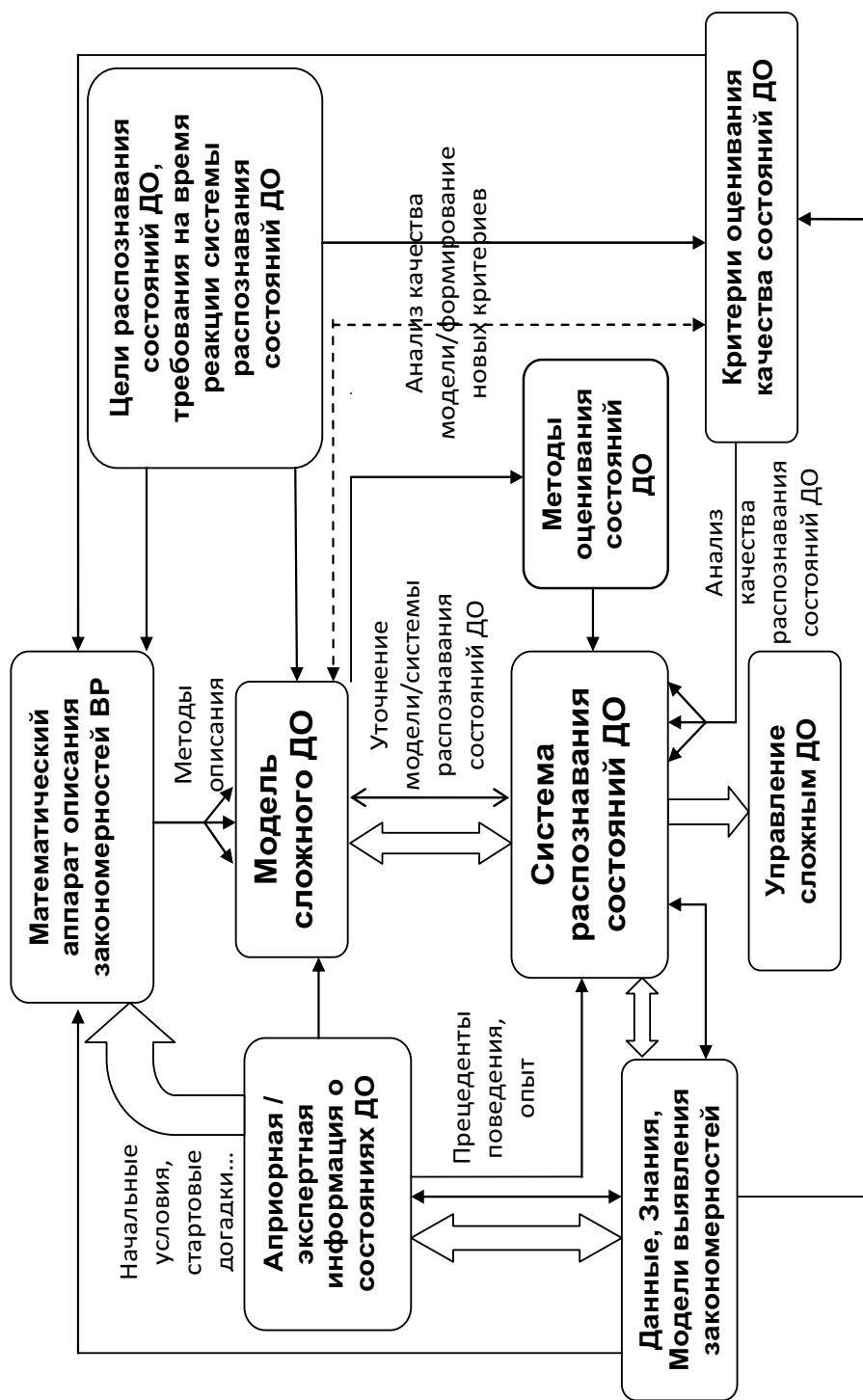


Рис. 0.1. Схема информационного и вычислительного взаимодействия моделирования динамического объекта и системы распознавания его состояний. Обозначения на рисунке: ВР – временной ряд; ДО – динамический объект; стрелки: широкие – поток данных (в т.ч. алгоритмов) и информации, тонкие – поток информации; пунктирные – не обязательность связи; тройные – указание на неоднозначность

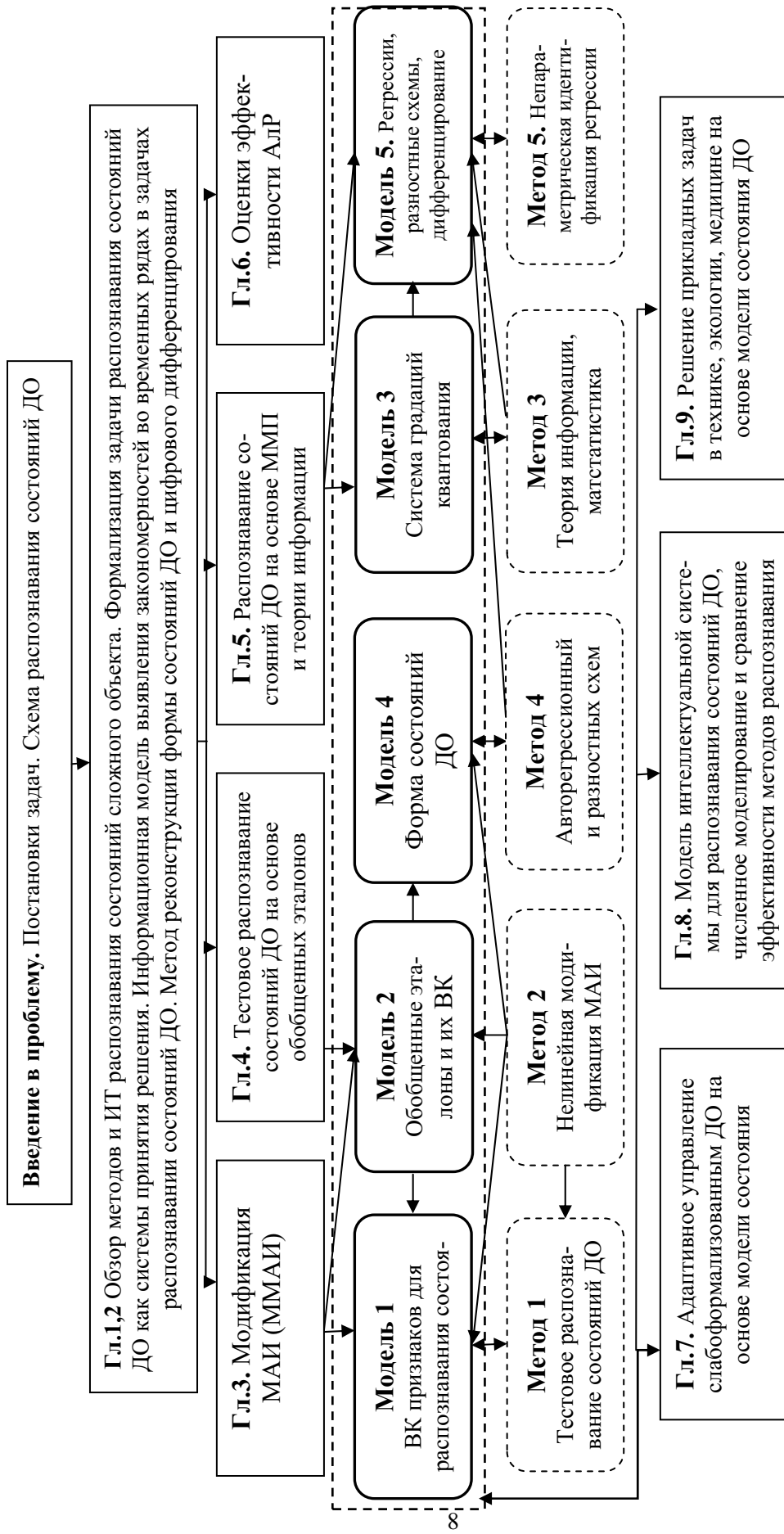


Рис.0.2. Структурная схема работы. Обозначения. ИТ - информационные технологии, ВК - весовые коэффициенты, ДО - динамический объект, ММП - метод максимального правдоподобия, МАИ - метод анализа иерархий, АЛР - алгоритмы распознавания состояний ДО

1. Задача адаптивного управления динамическим объектом в условиях неопределенностей с построением модели наблюдателя для неизмеряемых координат и неизвестных возмущений. Данная задача в общем аналитическом виде для нелинейных динамических объектов на настоящий момент не решена. В книге рассмотрен подход к решению этой проблемы на базе развития (гл. 7) метода АКАР (аналитическое конструирование агрегированных регуляторов¹), основанного на идеях самоорганизации - спонтанного перехода от неупорядоченного состояния к упорядоченному за счет совместного действия многих факторов²). Как показал И.Р. Пригожин³, переходы, перестройки в процессе самоорганизации неуправляемы (по крайней мере, в традиционном смысле), переход (выбор пути из возможных) совершается под действием неопределённых флуктуаций. Термины «неопределённые» процессы и «случайные» различают. Случайные процессы – это статистические, вероятностные процессы и явления, они обладают статистической устойчивостью, и к ним применим весь аппарат теории вероятностей (законы распределения, оценки, числовые характеристики и т.д.). И.Р. Пригожин показал, и это принципиальный момент, что в процессе переходных режимов (из состояния в состояние) закон больших чисел не работает.

2. Задача экологического диагностирования и прогнозирования.

3. Задача моделирования состояния здоровья пациента для оценки и прогноза его состояния.

Вторая и третья задачи относятся к так называемым слабоформализованным задачам, для которых построение точной математической модели вряд ли возможно и целесообразно. Оба эти объекта являются динамическими по своей природе и описываются рядом показателей (признаков), изменяющихся во времени, образуя многомерный стохастический нестационарный временной ряд с нелинейными трендами. Отметим также, что результаты регистрации показателей, как правило, процессы уникальные и «статистику» набрать либо невозможно, либо затруднительно.

4. Задача предсказания неизвестных значений атрибутов в базах данных связана с проблемой неспособности промышленных реляционных систем управления базами данных (СУБД) эффективно хранить и обрабатывать временные ряды, а именно этот класс СУБД обеспечивает основу крупных корпоративных информационных систем. Несмотря на то, что реляционные СУБД имеют все необходимые средства для описания временного ряда как одной из хранимых структур данных, низкая производительность обработки этой структуры, высокая трудоемкость описания прикладной логики, небольшая точность

¹ Синергетическую концепцию в теории управления поддерживал академик РАН А.А. Красовский, автором метода АКАР является А.А. Колесников.

² В США – это хаосодинамика, во Франции – наука о диссипативных структурах, а в России и Германии – это синергетика и нелинейная наука.

³ Известный бельгийский физико-химик, лауреат Нобелевской премии по химии 1977г.

предсказания (пропущенных, будущих) данных приводят к тому, что на практике СУБД не используются обработки временных рядов.

Поскольку регистрация изменений во времени любого технологического и природного процесса, по сути, есть временной ряд, то нетрудно продолжить список задач, где для описания ДО применяются временные ряды:

- задачи мониторинга, позволяющих в реальном времени следить за характером образования и развития нежелательных (потенциально опасных, катастрофических) состояний объекта;

- задачи построения вероятностной оценки состояния объекта или системы, находящейся в режиме управления;

- задачи управления в структурно сложных системах, являющихся нелинейными, многомерными и многосвязными, в которых протекают сложные переходные процессы и возникают критические и хаотические режимы и т.д.

Математический аппарат анализа временных рядов применяется, во-первых, для выявления закономерностей в поведении наблюдаемого процесса и прогнозирования; во-вторых, для выявления закономерностей в доступных наблюдениях реализаций случайного (и/или нелинейного) сопровождающего анализируемое явление процесса с целью мониторинга и управления. Несмотря на глубокую теоретическую проработку методов анализа временных рядов, их практическая реализация в настоящее время встречает серьезные трудности, связанные с многомерностью и нелинейностью реальных процессов.

Из приведенных примеров следует, что разработка методов, обладающих относительной простотой реализации, характеризующихся невысокими требованиями к ресурсам ЭВМ и при этом позволяющих строить достаточно точные модели многомерных нелинейных объектов и процессов, является весьма актуальной. Подмножеством таких методов являются эвристические методы – методы, не имеющие строгого теоретического обоснования, основанные на опыте и интуиции разработчика. Анализ условий применимости некоторых из этих методов и разработке новых посвящены исследования данной книги.

Замечания к изложению текста. Доказательства утверждений, теорем находится в приложениях с номерами, совпадающими с номерами глав.

Постановка задачи распознавания состояний электромеханических систем (ЭМС) с целью применения апостериорной информации для коррекции управления сложным объектом сделана В.Г. Букреевым; математический аппарат интеллектуального инструментального средства ИИС ИМСЛОГ на основе матричной модели данных и знаний разработан А.Е. Янковской, доказательство теорем, если это не оговорено специально, формализация постановки задачи распознавания состояний динамического объекта и подходов к ее решению с целью управления сложным динамическим объектом, осуществлены С.И. Колесниковой; остальные результаты получены в соавторстве.

«...большинство задач в нашем мире не описываются формулами, их не просчитаешь, в них неполная или крайне противоречивая информация» /Ю.И. Журавлев/

Глава 1. ОБЗОР МЕТОДОВ И ФОРМАЛИЗАЦИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ СОСТОЯНИЙ СЛОЖНОГО ДИНАМИЧЕСКОГО ОБЪЕКТА

В отношении реальных объектов (систем) термин «динамический объект» понимается в смысле: объект (система), описание которой возможно и целесообразно с помощью динамических моделей, причем не только детерминированных, но и стохастических, состояние которых изменяется (дискретно или непрерывно) во времени (см. обзор, например, в [29, 30]).

1.1. Обзор современных методов и информационных технологий распознавания состояний сложного динамического объекта

Краткий обзор, опирающийся на отечественные и зарубежные работы [17, 18, 25, 26, 28-35, 38, 43-45, 49-51, 57, 59, 64-68, 107, 109, 114-121, 127, 128, 137, 139, 147, 148, 150, 169] (в т.ч., обзорные [26, 30]), не претендуя на полноту, сделан, во-первых, с целью обоснования применимости (неприменимости) известных методов для решения задачи распознавания состояний ДО; во-вторых, с целью показать место обсуждаемых подходов в представляемой информационной модели (ИМ) к решению задачи распознавания состояний сложных ДО.

Реальные временные ряды (ВР), являясь дискретной моделью поведения динамических объектов, как правило, содержат параметрическую неопределенность, нестационарны и зашумлены. Обработка таких ВР требует разработки и использования различных методик и технологий, которые бы были свободны от предположений классических методов. В арсенал методов анализа динамики сложных динамических объектов входят статистические, вероятностные, логические, нечеткие и нейросетевые методы, методы нелинейной динамики и эвристические.

Статистические, вероятностные методы характеризуются относительной простотой реализации, однако требуют достаточно большого объема экспериментальных данных и не позволяют получать высокоточные модели многомерных объектов и процессов. По К. Грейнджеру⁴ [5] статистические методы, применяемые для стационарных временных рядов, способны показывать «абсолютно неверные результаты в случае их применения к динамическим рядам». Явление коинтеграции⁵, открытое им, означает, что характеристики определенных комбинаций ди-

⁴ Нобелевский лауреат 2003г.

⁵ Ряд, образованный линейной комбинацией нестационарных временных рядов и являющийся стационарным, называется коинтегрированным рядом (коинтеграционной регрессией).

намических временных рядов могут быть неизменными во времени, что позволяет корректировать статистические выводы.

Нейросетевые методы, логические методы и методы нечеткой логики позволяют строить более точные модели в условиях ограниченного набора обучающих данных, но при этом характеризуются относительной сложностью практической реализации и высокими требованиями к ресурсам ЭВМ.

Эвристические методы (не имеющие строгого теоретического обоснования, основанные на опыте и удачных прецедентах) обладают относительной простотой реализации, характеризуются относительно невысокими требованиями к ресурсам ЭВМ и при этом позволяющих строить достаточно точные модели многомерных нелинейных объектов и процессов.

Кратко перечислим условия применимости основных моделей и методов исследований ВР.

Регрессионные модели. Корректное применение регрессионных моделей предполагает постоянство средних величин (математического ожидания, дисперсии, ковариации). В связи с этим, для нестационарных ВР уточнение модели в случае зависимости указанных величин от времени может быть сделано следующими способами: 1) аналитическое моделирование такой зависимости; 2) переход к первым, вторым и т.д. разностям в нестационарных ВР, выражающих зависимость средних величин от времени (большое количество взятых разностей приводит к менее стабильным оценкам коэффициентов).

Автокорреляционные (авторегрессионные) модели. В моделях, использующих лаговую автокорреляцию (лаги - промежутки времени, на которых проявляется скрытая зависимость с.в.), используется выборочная автокорреляционная функция. Если функция модуля автокорреляционной функции имеет небольшое число максимумов и достаточно быстро спадает с ростом шага автокорреляции, то метод рекомендуется для анализа (стационарных) ВР. Для нестационарных ВР выяснение различия периодичности в корреллограммах, обусловленной физическими особенностями явления и внутренней взаимозависимостью между членами ряда, связано с усреднением корреллограммы по «окну» и переход к средней корреллограмме за период наблюдений, но размер периода усреднения («окна») подбирается «на ощупь».

Усложнение автокорреляционных моделей привело к адаптивным моделям, являющихся объединением двух схем: (взвешенных) скользящих средних и авторегрессии. Оценка текущего уровня определяется как взвешенное среднее всех предшествующих уровней, причем веса при наблюдениях убывают по мере удаления от последнего уровня. Вопрос об оптимальном выборе сглаживающих функций и весов решается субъективно, а на практике - переборным способом.

Методы спектрального анализа. Традиционно к этим методам относят использование преобразования Фурье (разложения сигнала на гармоники). В обобщенном смысле спектральным методом можно называть любое представ-

ление наблюдаемого сигнала в виде суперпозиции некоторых базисных функций, т.к. используется набор (спектр) функций-компонент (вейвлет-анализ).

Пример 1.1. В работе [110] качестве исходных признаков для определения состояния ДО принимались с.в. - величины значимых (превышающих заданный порог) амплитуд гармоник спектрограммы (периодограммы, выражающей зависимость суммы квадратов амплитуд синусоидальной и косинусоидальной гармоник от частоты гармоники) в разложении в ряд Фурье временного ряда. При решении задачи распознавания осуществлен переход от временных рядов к совокупности случайных величин - случайному вектору. В зависимости от состояния исследуемого ДО математическое ожидание исходного вектора признаков предполагалось различным. Наблюдаемые его значения, полученные на основании измерений ВР в различные интервалы времени, интерпретировались как векторы в евклидовом пространстве. Осуществив предварительные эксперименты при заранее известных состояниях ДО, получали некоторые группы векторов, которые будут соответствовать этим состояниям (обучающую выборку). Наиболее простая процедура распознавания была основана на нахождении евклидовых расстояний между исходным вектором признаков для неизвестного состояния динамической системы и выборочными средними векторов групп. Выбиралась группа, к которой это расстояние оказывалось наименьшим.

Укажем недостатки данного подхода: 1) 1-й вариант процедуры распознавания годен при условии различия выборочных средних векторов групп, соответствующих разным состояниям ДО, что для ряда технических процессов мало приемлемо (длительности состояний ДО не только не являются постоянной величиной, но и смена состояний ДО не происходит скачкообразно), к тому же разная «частотная картина» разных состояний может приводить к одинаковым средним значениям; 2) выбор признаков (частоты, которым соответствуют максимумы спектрограммы, а также соответствующие им значения периодограмм при разложении фрагментов ВР в ряд Фурье) при построении процедуры распознавания корректен в предположении стационарности ВР на отрезке времени, соотнесенном состояниям ДО (непериодический СП нельзя представить в виде ряда Фурье с взаимно независимыми коэффициентами).

Спектральное разложение сигнала имеет особенность «размывания» спектральных составляющих (напр., [57]) из-за конечной длительности обрабатываемых наборов данных (обучающей выборки). Поскольку периодических функций в природе не существует, то пользоваться спектральными методами рекомендуется в следующих случаях:

- интервал наблюдений содержит большое число колебаний (в теории сигналов, $|t_1 - t_0| > 30T$) с наибольшим периодом T , или «длинные» ряды, когда краевыми эффектами можно пренебречь;
- форма колебаний на каждом периоде неизменна на интервале наблюдения;

- наибольший период T и периоды гармоник постоянны на интервале наблюдения.

Так, для временных рядов технического (например, ЭМС), экономического происхождения все условия не выполняются, поэтому попытки использования спектральных методов в этой предметной области (К. Грейнджер [5]) приводят к неадекватным выводам.

Кроме того, доказано, что при неизвестной природе сигнала (отсутствует математическое описание шума, как и в данной работе) нет оптимального спектрального окна [57].

Вейвлет-анализ полезен для исследования нестационарных сигналов, содержащих сегменты с различным характером поведения, для существенно неоднородных сигналов (импульсного типа) с особенностями (разрывами, изломами, разрывами производных более высокого порядка), поскольку позволяет локализовать особенности и выявить их характер. Базис состоит из функций (вейвлетов), локализованных по времени и по частоте. Анализ ВР на основе данной группы методов существенно осложняется нелинейностью динамических характеристик рассматриваемой динамической системы.

Бутстреп-метод [47, 169] состоит в том, что по имеющимся N наблюдениям за случайной величиной, образующим по предположению выборку из генеральной совокупности, строится выборочная функция распределения, из которой извлекаются выборки с возвращением того же объема N с равными вероятностями извлечения каждого значения. Однако, в силу нестационарности, размноженные выборки, полученные с помощью выборочной функции распределения на каком-либо отрезке, будут отличаться от выборки на основе истинной функции распределения, сильно меняющейся во времени.

Нейросетевой метод. Целью метода при нахождении корреляционной связи между элементами временного ряда [161] является отыскание «причины» наступления того или иного события и определение вероятности последнего [166] на основе эмпирической вероятности наступления похожих случаев в период обучения на достаточно большом количестве выборок некоторого объема. К недостаткам метода относят относительность выдаваемых ответов, высокую вычислительную трудоемкость, отсутствие теоретических оценок гарантий применения метода. Трудоемкость обучения нейронных сетей, немалая даже в случае стационарности ВР, в условиях нестационарности ВР только возрастает (более чем в 2 раза) с одновременным существенным уменьшением обобщающей способности [47, 44, 158]. При возрастании искажений во времени, нейросети хуже справляются с задачей распознавания (число распознанных объектов резко падает).

Методы нелинейной динамики исследования сложных стохастических динамических систем систематизированы относительно двух подходов: безмодельные исследования или непосредственный анализ, куда относят спектральный анализ (разложение в ряды Фурье, Карунена–Лоэва, по базовым вейвлет-функциям), полу-

чение некоторых числовых (средние значения, корреляционные размерности, показатели Ляпунова и т. д.) и моделирование ДО (реконструкция).

Обоснование применения методов нелинейной динамики⁶ к анализу временных рядов (и обратно) состоит в следующем. Предельная (бесконечное время) структура нелинейного хаотического объекта (аттрактор, подмножество фазового пространства, притягивающее траектории в пределе), может быть восстановлена через измерение только одной наблюдаемой этого динамического объекта, фиксированной как временной ряд.

Динамический объект может быть описан, например, обыкновенными дифференциальными уравнениями $\dot{x} = F(x(t))$, или, если время дискретно $t = k\Delta$, уравнениями вида: $x_{k+1} = f(x_k)$. Эволюция динамического объекта описывается с течением времени в некотором фазовом пространстве $X \subset R^d$.

Временные ряды возникают как последовательность наблюдений $\{y_k, y_k = s(x_k), k \geq 1\}$, полученных с помощью некоторой функции измерения $s(\cdot)$. Последовательность $\{y_k\}$ не порождает многомерное фазовое пространство ДО, но на основе технического приема, согласно методу Грассбергера и Прокаччи⁷, возможно раскрытие многомерной структуры ДО, используя имеющиеся данные только одной составляющей многомерного временного ряда.

Процедура реконструкции фазового пространства и восстановления аттрактора системы при динамическом анализе временного ряда, сводится к построению пространства вложений с помощью метода задержки. Векторы в новом пространстве вложения сформированы из значений временного ряда скалярных измерений с временным запаздыванием: $\bar{y}_k = (y_{k-(p-1)\tau}, y_{k-(p-2)\tau}, \dots, y_k)$.

Число элементов p называется размерностью вложения, интервал времени τ обычно называется задержкой (лагом), при числе n измерений во временном ряду число вложенных векторов равно $n - (p-1)\tau$. В теоремах Такенса и Сойера [18] утверждается, что при подходящем подборе размерности вложения p (достаточно большом), реальный аттрактор динамического объекта и «псевдоаттрактор», восстановленный в лаговом пространстве по временному ряду, являются топологически эквивалентными и обладают одинаковыми обобщенными фрактальными (корреляционными) размерностями, ляпуновскими показателями и другими численными характеристиками.

В том случае, если анализируемый временной ряд является реализацией случайного процесса, то восстановленный аттрактор представляет собой бесструктурное множество точек, которое при последовательном бесконечном увеличении размерности вложения пространства p равномерно его заполняет (кор-

⁶ Всякий объект с хаотическим поведением, как правило, имеет нелинейную модель описания, но не всякий «нелинейный» объект имеет хаотическое поведение.

⁷ Grassberger P., Procaccia I. Measuring the strangeness of strange attractors. // Physica D 9. (1983).

реляционная размерность стремится к бесконечности). Однако невозможно установить, является ли временной ряд реализацией случайного процесса или последовательностью измерений фазовой траектории сложного хаотического объекта по конечному отрезку временного ряда. Кроме того показано⁸, нелинейные стохастические модели типа ARCH могут генерировать динамику, схожую с хаотическими отображениями, и тест на хаос на основе корреляционной размерности может указывать наличие именно такого вида нелинейности.

Метод сингулярного спектрального анализа (метод главных компонент) разработан в теории нелинейных динамических систем для исследования автокорреляционной зависимости. В методе рассматривают выборку некоторого объема n и выбирают скользящее окно объема L . Выборки $(x_1, \dots, x_L), (x_2, \dots, x_{L+1}), \dots$ рассматриваются как векторы L -мерного линейного пространства. Оформи́в выборки в виде матрицы размером $L \times (n-L+1)$, находят ранг этой матрицы, определяют базис (главные компоненты ряда) в L -мерном пространстве, связывают значение x_{k+1} с базисными векторами $\mathbf{y}(k, L)$, $\mathbf{y}(k, L) = (x_k, x_{k-1}, \dots, x_{k-L+1})$ используя некоторую подходящую функцию: $x_{k+1} = f(\mathbf{y}(k, L))$. Такая динамическая система порождает стационарный временной ряд, называемый главной частью анализируемого. Для этой системы можно определить траекторию, аттрактор и его размерность, функцию распределения точек траектории и другие характеристики.

Вопрос о размерности самой матрицы и количественной зависимости от этой размерности числа базисных векторов остается в этом методе открытым. В [65], например, предлагается использовать «графический критерий близости выделяемых компонент идеальному гармоническому сигналу» для определения подходящей длины выборки.

Метод анализа временных рядов, основанный на алгебраическом подходе [139], решает задачу выделения и классификации траекторий аномального поведения ДО на сопровождающих его функционирование временных рядах. Сложность задачи заключается в том, что закономерности (образы состояний ДО), выявленные на разных реализациях, могут существенно отличаться, что может приводить к искажениям выявленных особенностей (признаков) по времени и по амплитуде, и, следовательно, существенно ограничивать применимость большинства традиционных алгоритмов распознавания.

В основе алгебраического подхода лежит идея разметки точек ряда аксиомами (условиями). На первом этапе (обучения) строится система аксиом, а на втором (этапе распознавания) - каждому классу «специфичного» поведения ставится в соответствие «эталонная» последовательность аксиом. На этапе распознавания предъявленный ряд (фрагмент) размечается выбранной системой акси-

⁸ J.A. Sheinkman, B. LeBaron. Nonlinear Dynamics and Stock Returns. Journal of Business 62, 1989.

ом, и на полученном ряде аксиом осуществляется поиск подпоследовательностей, близких к эталонным. Ключевой проблемой при этом является проблема выбора системы аксиом для данной конкретной задачи, решаемая, как правило, ручной подборкой, что является большой проблемой для многомерных временных рядов. В данной работе будет показан подход к частичному решению данной задачи для стохастических ВР (п. 2.4).

Известны работы, где используется либо генетический алгоритм для решения задачи автоматического выбора подходящей системы аксиом, реализующий введение трех классов операций скрещивания и мутации систем аксиом и рассмотрение нескольких видов селекции, либо алгоритм на основе совокупностей нейросетей [44].

Метод эмпирической модовой декомпозиции (EMD) на основе преобразования Гильберта-Хуанга⁹ - новый метод анализа нелинейных и нестационарных данных, представляющий собой метод декомпозиции данных с извлечением внутренних функций данных и преобразованием в локальные энергетические и мгновенные частоты. Особенностью алгоритма на основе данного метода является отсутствие ограничений на частотно-временное разрешение, полная адаптивность разложения нелинейного нестационарного сигнала на набор монокомпонентных составляющих, каждая из которых описывает собственный процесс, с последующим анализом частотно-временного распределения составляющих на основе преобразования Гильберта.

Данный метод свободен от физической сущности обрабатываемых данных и статистики распределения шумовых сигналов и дестабилизирующих флуктуаций, может эффективно использоваться для очистки от шумов данных, не критичных во времени обработки, так как существенно использует интерполяционное сглаживание (сплайнами).

Пакеты программ статистической обработки данных (численные методы в том числе). Пакеты STATISTICA¹⁰, SAS, SPSS и др. включают анализ ВР, дисперсионный, линейный и нелинейный регрессионный анализ; анализ прерванных ВР (т.н. рядов с интервенциями); реализацию моделей экспоненциального сглаживания, спектрального анализа одного ряда и кросс-спектрального анализа двух рядов. Алгоритмы рассчитаны на аппроксимацию имеющихся данных, причем данных стационарных рядов, поскольку подбор скользящего окна остается за пределами возможностей пакетов, а методы автоматизированного «подбора» модельных трендовых функций отсутствуют. Выбор наиболее адекватной модели стационарных ВР проводится по известным алгоритмам спецификации моделей [25, 28, 33, 107], позволяющим отобрать оптимальное число

⁹ Huang N.E., Shen Z., Long S., et al. Proc. R. Soc. London A. - 454. - 1998. - pp. 903-995.

¹⁰ Универсальный пакет, от которого другие пакеты отличаются включением или, напротив, не включением более или менее стандартных тестов и критериев, которые содержатся в специализированных справочниках.

параметров в рамках дисперсионного и корреляционного анализов, однако для нестационарных ВР аналитически обоснованные алгоритмы отсутствуют.

К особенностям и условиям применения метода Бокса-Дженкинса ARIMA для анализа нестационарных рядов относятся следующие:

1) не предполагается задания математической модели из физической природы данной временной серии. Алгоритм, подстраивая внутренние параметры, выбирает наиболее подходящую модель прогнозирования из заданного общего класса моделей;

2) практическая ценность модели высокого порядка, в результате которой остатки оказываются чистым белым шумом, невелика, так как применение такой модели может показывать разброс 95%-го прогноза в 2-4 и более раза;

3) неприменимость для технических систем реального времени с критичным временем реакции.

Общий недостаток статистических моделей (в том числе, инструментов прогнозирования нестационарных временных рядов ARIMA и GARCH) заключается в том, что они используют исторические данные, не учитывая физические особенности анализируемого ряда. В связи с этим резкие изменения условий будут учтены только через определенный промежуток времени, до момента его достижения предсказания будут некорректны.

Метод выборочного оценивания функции распределения. При определении ошибки оценивания нестационарного временного ряда надо учесть факторы: конечность выборки и различие распределений для разных выборок вследствие нестационарности процесса. Разные методы имеют неодинаковую чувствительность точности аппроксимации данных к действию указанных факторов. Для нестационарного ВР теоремы об эффективности, состоятельности и асимптотической нормальности выборочных оценок и их дисперсий в общем случае не выполняются.

Методы технической диагностики решают задачи распознавания состояний системы, определения причин нарушения работоспособности, а также установления вида и места дефекта. Методы условно классифицируют [50] на три большие группы: методы параметрической идентификации, статистические и вероятностные методы, логические методы.

Методы параметрической идентификации строятся на основе минимизации ошибки между измерениями выходного сигнала ДО и выходом модели ДО при условии подачи на вход сигнала определенного типа. Различают, при этом, подходы к идентификации параметров: 1) на основе применения нелинейного МНК и ММП [64, 67]; 2) на основе использования импульсных характеристик [135]; 3) на основе метода множителей Лагранжа [32]. Однако вопрос о типе структуры модели и ее минимальной сложности не решается.

Статистические и вероятностные методы. Одна группа методов основана на решении интегрального уравнения Винера-Хопфа при различных ограничениях; другая - для определения диагностической характеристики объекта

требует решения уравнения Риккати с заданными начальными условиями; третья - основана на методах последовательного анализа Вальда и условных марковских процессов, развитых в работах Р.Л. Стратоновича, А.Н. Ширяева, Э.М. Хазен и др., и позволяет получить двухэтапную процедуру диагностики динамической системы, реализация которой вычислительно возможна. На 1-м этапе определяется оптимальная проверка, на 2-м проводится оценка диагностической характеристики элемента проверки, по значениям которой уточняются априорные сведения о состоянии блоков, включенных в элемент проверки. В целом группа методов характеризуется относительной простотой реализации, обладает хорошей обобщающей способностью при невысоких требованиях к вычислительным ресурсам со всеми особенностями, присущих данным методам.

Логические методы основаны на детерминированной двоичной оценке состояния блоков рассматриваемого объекта и выборе признаков неисправностей по таблице состояний, задающей отображение пространства параметров на пространство наблюдений Y . В работе [109] изложена строгая математическая постановка задачи поиска неисправности и контроля работоспособности и получены оптимальные и квазиоптимальные стратегии поиска неисправного блока в случае пересекающихся проверок различной стоимости на основе критерия минимума средних потерь. Однако, во-первых, построение логических моделей возможно только тогда, когда связь между пространством параметров и пространством наблюдений определена однозначно (детерминированная связь между векторами указанных пространств нарушается из-за действия помех и влияния различных случайных факторов как на объект, так и на систему регистрации записей динамики ДО, являющуюся основным источником информации); во-вторых, весьма ограничительным условием к применению логических методов относятся вычислительные сложности алгоритмизации выбора признаков неисправностей по таблице состояний; требование наличия некоторых трудно/недоступных статистических данных.

Перечисленные методы имеют большой процент ошибок при использовании в задачах обнаружения предвестников зарождающихся дефектов (разрушений) различных узлов, связанный с тем, что разброс величин измеряемых параметров даже в группе одинаковых бездефектных машин, как правило, превышает изменения, характерные для появления зарождающихся дефектов.

Отметим, что общего формализованного подхода к решению проблемы распознавания состояний ДО на основе существующих подходов к анализу ВР, являющихся дискретной моделью поведения сложных (нелинейных) ДО, не существует. Отметим также, что создание полной модели для сложного объекта (системы) вряд ли целесообразно, так как в силу теоремы Тьюринга, такая модель будет столь же сложной, как и объект.

1.2. Основные понятия и определения

В силу терминологического разночтения остановимся кратко на далее используемых понятиях, приведенных описательно.

Под сложным динамическим объектом будем понимать объект, характеризующийся свойствами: 1) отсутствие полного аналитического описания (недостаточность априорной информации [130] для построения его адекватной модели); 2) нелинейность имеющихся моделей описания и/или нестационарность сопровождающего его поведение случайного процесса, (слабопредсказуемость). Далее будем также использовать термин «слабоформализованный динамический объект». ДО является нелинейным, если описываемые его дифференциальные уравнения нелинейные, или ВР, описывающий поведение его характеристик, имеет нелинейный тренд.

В книге будут рассмотрены следующие модели ДО: случайный процесс с дискретным временем, множество выборочных значений которого образуют временной ряд; система (в общем виде, нелинейных) дифференциальных уравнений; система разностных уравнений.

Определение 1.1. Пусть дано вероятностное пространство (Ω, \mathcal{F}, P) . Случайный процесс (СП) (случайная функция времени) определяется как параметризованное семейство $\{X_t\}_{t \in T}$ случайных величин $X_t(\cdot): \Omega \rightarrow \mathbb{R}$, $t \in T \in \mathbb{R}$, где T - параметрическое множество. Если T - дискретное множество (например, $T = \{0, 1, 2, \dots\}$), то $\{X_t\}_{t \in T}$ - СП с дискретным временем. При фиксированном $\omega = \omega_0 \in \Omega$ величина $x_t = \xi(t, \omega_0)$ как функция параметра t называется реализацией СП (траекторией). Значение x_t при фиксированном $t = t_k$ называется выборочным значением СП $x_k = x(t_k)$.

Определение 1.2. Временной ряд – множество выборочных значений случайного процесса $\{x_k = x(t_k), t_k \in T\}$. Будем полагать в работе: $T = \{0, \Delta, 2\Delta, \dots\}$, или $t_k = k\Delta$, $\Delta > 0$.

Состоянию $x(t)$ в некоторый момент t в фазовом пространстве соответствует точка (x_1, \dots, x_n) – изображающая точка (она изображает мгновенное состояние). Изменение состояния системы во времени будет соответствовать движению изображающей точки в фазовом пространстве по определенной траектории, которая называется фазовой. Совокупность характерных фазовых траекторий называют фазовым портретом системы.

Под состоянием сложного ДО будем понимать подмножество фазового пространства ДО, фрагмент фазовой траектории (сегмент ВР), обладающих определенными свойствами (ограничениями); обобщенное образное описание особенностей фазовой траектории ДО по группе параметров (признаков), задаваемых на основе экспертного анализа.

Под объектом будем понимать фрагмент (сегмент, отрезок) временного ряда, характеризующихся рядом показателей (признаков).

Класс – группа объектов (или образов по [156]), отобранных по одному или нескольким показателям (признакам).

Под образом (состояний ДО) будем понимать совокупность характеристических признаков, отличающую объект одного класса-состояния от объекта другого класса-состояния. В подходе к автоматизации численного исследования ДО на основе использования методов распознавания образов и статистического моделирования [121] образы – это фазовые траектории, установившиеся движения и их области притяжения.

Часто используемое понятие «информативность» не является в научной литературе однозначно формализованным. В этой работе под информативностью признаков будем понимать некоторую информационную характеристику, описывающую совокупность признаков, например, на основе информационных критериев (Шеннона, Кульбака), на основе весовых коэффициентов признаков, характеризующих долю «полезной» информации в общем объеме признаков (в смысле их способности различать объекты из разных образов).

Первичный признак – это характеристика объекта, которую можно измерить или вычислить на основе измерений (например, значение ВР, значение скользящего среднего).

Характеристический признак (далее, признак) - характеристика объекта, возможно полученная скаляризацией первичных признаков (среднее значение, эталон, полученный по определенному правилу, частота проявления первичных признаков и т.д.). Число характеристических признаков, как правило, существенно меньше числа первичных признаков, а информативность - выше.

Классификационный признак (также решающее правило) - признак, определяющий механизм (метод) разбиения объектов на классы эквивалентности. В теории многокритериального анализа метод структурирования множества альтернатив (гл. 3) также принято называть решающим правилом.

Под обучающей выборкой будем понимать множество конечных многомерных временных рядов разной размерности и длительности (массив конечных отрезков траекторий или реализаций случайного процесса как модели ДО), являющихся результатом реальных измерений (модельных значений), относительно которых известно их «происхождение» (номер, метка класса).

Задача эмпирического предсказания заключается в следующем: получив выборку данных, необходимо предсказать определённые свойства аналогичных данных, которые станут известны позже, и оценить точность предсказания.

Метод минимизации эмпирического риска в статистической теории восстановления зависимостей [42] предполагает, что множество объектов X является вероятностным пространством (сильная вероятностная аксиоматика), и все рассматриваемые выборки являются случайными, независимыми, одинаково

распределенными. В слабой вероятностной аксиоматике [47] $\mathbf{X}=\mathbf{X}^L$ – конечное множество попарно различных объектов.

Определение 1.3. Пусть задано множество попарно различных объектов $\mathbf{X}=\mathbf{X}^L$. Методом обучения называется отображение μ , которое произвольной конечной обучающей выборке $X\subset\mathbf{X}$, ставит в соответствие определенный алгоритм $a=\mu(X)$, выбирая его из некоторого семейства алгоритмов A . Говорят, что метод μ строит алгоритм a по обучающей выборке X .

Определение 1.4. Пусть существует бинарная функция $\chi: A\times\mathbf{X}\rightarrow\{0,1\}$, называемая индикатором ошибки. Если $\chi(a,x)=1$, то говорят, что алгоритм a допускает ошибку на объекте x .

Качество алгоритма a на конечной выборке U характеризуется либо вероятностью ошибок, либо частотой ошибок $v(a,U)=|U|^{-1}\sum_{x\in U}\chi(a,x)$.

Качество алгоритма a на отдельном разбиении характеризуется частотой ошибок на обучении $v(a,X)$ и на контроле $v(a,\bar{X})$, $\bar{X}=\mathbf{X}\setminus X$. Поскольку вычисление этой величины не представляется возможным, то ставится задача определения условий, при которых она «близка» к эмпирическому риску $v(a,X)$.

Процесс обучения состоит в построении алгоритма $a^*\in A$ путем минимизации эмпирического риска: $a^*=\arg\min_{a\in A}v(a,X)$, $X\subset\mathbf{X}$. Достаточным условием

обучаемости семейства алгоритмов является малое значение функционала:

$$P_\varepsilon(A)=\mathbf{P}\left\{\sup_{a\in A}(v(a,\bar{X})-v(a,X))>\varepsilon\right\}, 0\leq\varepsilon\leq 1, \text{ равномерного отклонения частоты}$$

ошибок в двух выборках. В силу неизвестности получаемого в результате обучения алгоритма a оценивается максимальное отклонение, достигаемое на наилучшем алгоритме.

Определение 1.5. Алгоритм a называется корректным на выборке X , если $v(a,X)=0$. Метод μ называется корректным на выборке X , если алгоритм $\mu(X)$ корректен на X .

Определение 1.6. Функцией роста $\Delta^A(L)$ семейства A называется максимальное количество различных бинарных векторов вида $\{\chi(x_i,a(x_i))\}_{i=1}^L$, порождаемых всевозможными алгоритмами $a\in A$ на произвольной выборке \mathbf{X}^L .

Определение 1.7. Минимальное число h , при котором $\Delta^A(h)<2^h$, называется емкостью семейства алгоритмов A . Если такого числа h не существует, то говорят, что емкость семейства бесконечна.

Пример 1.2. Линейный классификатор разбивает точки на плоскости на два класса прямой линией. Максимальное число (не лежащих на одной прямой) точек, которое может быть разделено прямой линией на два класса всеми возможными способами равно трем ($2^3=8$); множество из четырёх и более точек - всеми возможными способами ($2^4=16$) разделить одной прямой уже нельзя; ём-

кость семейства линейных классификаторов $h=3$. Алгоритм ближайших соседей имеет бесконечную ёмкость.

Одним из основных критериев оценки качества алгоритмов распознавания состояний ДО является критерий скользящего контроля.

Процедура скользящего контроля заключается в следующем. Производится M разбиений выборки \mathbf{X}^L на обучающую подвыборку длины ℓ и контрольную длины k , $\mathbf{X}^L = X_g^l \cup X_g^k$, $L = l + k$, $g = \overline{1, M}$. Оценка скользящего контроля для

функции $\xi : \{1, \dots, M\} \rightarrow \mathbb{R}$ определяется как среднее $\hat{\xi} = \frac{1}{M} \sum_{g=1}^M \xi(g)$. Разбиения

строятся по стандартной методике $m \times q$ -fold cross-validation: генерируется m случайных разбиений выборки \mathbf{X}^L на q блоков примерно равной длины, каждый блок поочерёдно становится контрольной выборкой ($M = mq$).

Определение 1.8. Модель, минимизирующая среднюю ошибку на контрольной подвыборке \bar{X} по всем разбиениям исходной выборки на обучающую X и контрольную \bar{X} ($\mathbf{X} = X \cup \bar{X}$), называется оптимальной по критерию скользящего контроля.

Явление получения частоты ошибки на обучающей выборке (как оценки вероятности ошибки) существенно более низкой, чем на неизвестной выборке, называют переобучением.

Способность алгоритмов восстанавливать неизвестную зависимость по конечной выборке данных называют обобщающей способностью.

График зависимости $P_\varepsilon(A)$ от ℓ при фиксированном k используют и для оценивания достаточной длины обучающей выборки для данного метода в данной задаче и для оценивания риска переобучения.

Оценки статистической теории [42] выводились при условии, что распределение вероятностей на множестве объектов существует, но неизвестно. Если семейство A имеет конечную ёмкость h , то в сильной вероятностной аксиоматике имеет место равномерная сходимост частот: $P_\varepsilon(A) \xrightarrow{l \rightarrow \infty} 0$. Показано [42], что

оценка вероятности ошибки по критерию скользящего контроля является несмещенной и эффективной. Однако, данную оценку на практике применять весьма затруднительно, т.к. во-первых, оценивание ёмкости является отдельной непростой задачей для каждого конкретного семейства; во-вторых, оценка $P_\varepsilon(A) \leq \Delta^A(2l)1.5 \cdot \exp(-\varepsilon^2 l)$ становится тривиальной при бесконечной ёмкости, присущей большинству семейств корректных алгоритмов, гарантирующих отсутствие ошибок на обучающей выборке.

В работах К.В. Воронцова [47] в слабой вероятностной аксиоматике (все перестановки конечного множества попарно различных объектов, называемого генеральной выборкой \mathbf{X} , имеют равные шансы реализоваться), получены верх-

ние оценки комбинаторных функционалов, аналогичные оценкам вероятности равномерного отклонения частоты ошибок в двух подвыборках.

На основе полученных оценок утверждается, что функционалы скользящего контроля (комбинаторные оценки качества): 1) характеризуют обобщающую способность алгоритмов не хуже вероятностных; 2) справедливы для любой выборки, не обязательно случайной, независимой и одинаково распределенной; 3) зависят от сложности не всего семейства алгоритмов, а от локального его подмножества, реально получаемого в результате обучения. Обоснованы избыточность принципа равномерной сходимости, вторичность конструкции семейства алгоритмов по отношению к методу обучения и «принцип соответствия» статистической теории и комбинаторной для задач анализа данных.

Под задачей распознавания в данной работе будем понимать получение информации о типе (виде) состояния динамического процесса на основе конечного интервала наблюдений значений выходных характеристик (показателей). Под системой распознавания реального времени будем понимать систему, гарантирующую заданное время принятия решения о классе измеряемого объекта.

Под дискретизацией реализации СП $Y(t)$ понимается разбиение непрерывного процесса по временной составляющей (на графике - по горизонтали); под квантованием реализации СП понимается разбиение диапазона значений (на графике - по вертикали) сигнала на отрезки равной длины. Под градациями квантования будем понимать уникальные значения уровней квантования. Градации состоят из одного значения при стандартном квантовании, из нескольких значений (мета-градации) при нестандартном квантовании.

Определение 1.9. Мультимножество $G = \{k_G(x) \bullet x \mid x \in U, k_G(x) \in Z_+\}$ [131] - совокупность пар значений $k_G(x) \bullet x$ элементов $x \in U$ (U - домен) и кратностей вхождения элемента $x \in U$, определяемых функциями $k_G: U \rightarrow Z_+ = \{0, 1, 2, \dots\}$, что обозначено символом « \bullet ». Мультимножество $G = (g_1, g_2, \dots, g_n)$ можно задать его основанием $S(G) = (x_1, x_2, \dots, x_n)$ - кортежем всех разных элементов из U и кортежем кратностей $[G] = (k_1, k_2, \dots, k_n)$ элементов основания мультимножества, называемым первичной спецификацией мультимножества. Если мультимножество $G(\omega)$ содержит элементы, являющиеся случайными числами, то $G(\omega)$ будем называть стохастическим (случайным) мультимножеством (СММ).

Под закономерностью будем понимать: а) подобласть признакового пространства, содержащую объекты одного класса (полная закономерность) или преимущественно одного класса (частичная закономерность); б) подмножества характеристических признаков, влияющими на различимость объектов из разных образов, устойчиво наблюдаемыми для объектов из обучающей выборки и проявляющимися на других объектах той же природы; в) весовые коэффициен-

ты характеристических признаков, отражающие их индивидуальный вклад в различимость объектов из разных образов [175].

1.3. Формализация задачи распознавания состояний динамического объекта

Модель распознавания состояний сложного динамического объекта (рис.1.1), будем строить, опираясь на классические варианты идейного содержания систем принятия решений [39, 40], формируя следующие уровни указанной проблемы:

- 1) цели, на котором будут определяться целевые задачи, анализу которых будет подчинена разрабатываемая модель, а также терминология и тезаурус;
- 2) частных постановок задач, решение которых возможно на основе модели;
- 3) математических и эвристических методов, применяемых для распознавания состояний ДО, формальных моделей для принятия многокритериальных решений;
- 4) алгоритмов, реализующих используемые методы и подходы к распознаванию состояний ДО;
- 5) выбора «наилучшего» алгоритма (совокупности алгоритмов) для решаемой задачи;
- 6) программной реализации алгоритмов;
- 7) критериев, ограничений, априорной информации.

Уровень цели. Цель исследований - повышение оперативности, достоверности и полноты предварительного анализа временных рядов с выявлением структурных и динамических особенностей сложных (слабоформализованных) объектов на основе создания комплекса методов и реализующих их алгоритмов выявления закономерностей характерных областей фазового пространства ДО (фазовых траекторий), называемых состояниями ДО.

Общая постановка задачи. Пусть задан случайный процесс (СП)

$$\{X, Y\} = \{X(t), Y(t), t_0 \leq t \leq T\}, Y(t) = f(X(t), \xi(t)), \quad (1.1)$$

характеризующий состояние ДО, где $X(t)$ – вектор ненаблюдаемых переменных состояния системы; $Y(t)$ – случайная наблюдаемая векторная функция; $\xi(t)$ – неизвестный шум с ограниченной дисперсией. Относительно динамики поведения СП на $[t_0, T]$ выдвинуто $I > 0$ альтернативных гипотез $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_I\}$, составляющих полную группу событий и интерпретируемых как классы состояний ДО. Другими словами, временной фрагмент реализации СП может находиться только в одной из заданных (экспертом, или выделенных по обучающей выборке и косвенным признакам) областей (классов) Ω_i ,

$i = \overline{1, I}$. Наблюдение осуществляется в соответствии с дискретным планом $t \in \{t_0, t_1, \dots, t_n\}$, $t_j = t_0 + j\Delta$, $j = \overline{0, n}$. Задача состоит в отнесении наблюдаемых в момент t или на некотором фиксированном интервале наблюдаемых значений реализации $Y(t)$ к одному из классов Ω_i , $i = \overline{1, I}$.

Уровень частных постановок задач. Приведем примеры содержательных задач прикладного характера, интерпретирующих общую задачу (1.1).

Задача 1. Мониторинг качества выпускаемой продукции по исходным данным в виде временного ряда (во многих областях, несмотря на АСУ, фактический контроль за качеством осуществляют эксперты на основе их личного опыта и интуиции).

Задача 2. Алгоритмизация распознавания состояний экологического объекта, модель характеристик которого задана в виде многомерного временного ряда. Прогнозирование и оценивание скорости изменения качества природных ресурсов и объектов с целью своевременного обнаружения явлений, приводящих к необратимым последствиям (предвестников).

Задача 3. Определение метки (номера) состояния ДО по экспериментальным показаниям датчиков в реальном времени на основе правил, сформированных на обучающей выборке - экспериментальных данных показаний периодически опрашиваемых датчиков при разных состояниях ДО в виде набора ВР, характеризующих развитие процесса.

Задача 4. Пусть задан непрерывный оператор, который взаимно однозначно отображает элементы $f(t, \theta)$ одного метрического пространства X_1 в элементы $y(t)$ другого метрического пространства X_2 (операторное уравнение):

$$Af(t, \theta) = y(t) + \xi(t) \quad (1.2)$$

где θ - вектор параметров модели, A - оператор отображения функциональной зависимости, ξ - вектор ошибок наблюдений и моделирования.

Например, в модели $\sum_{i=0}^l \theta_i x^{(i)}(t) = f(t)$, где $x^{(i)}(t)$ - производная i -го порядка выходного сигнала линейной системы, t - параметр времени, различным значениям входной величины $f(t)$ соответствуют выходные величины $x(t)$.

В прямой задаче по заданной $f(t, \theta)$ требуется найти $y(t)$ по наблюдениям $z(t) = y(t) + \xi(t)$; в обратной задаче требуется найти решение операторного уравнения (1.2) в заданном классе функций $f(t, \theta)$ по измерениям z_1, z_2, \dots, z_m функции $y(t)$ в m точках t_1, t_2, \dots, t_m .

Решение уравнения (1.2) является некорректным в том смысле, что сколь угодно малым изменениям наблюдаемой функции y могут соответствовать сколь угодно большие изменения зависимости f ; двум различным моделям пока-

зателей f_1 и f_2 , искаженных шумом, будут соответствовать очень похожие измерения y_1 и y_2 (и наоборот).

К существенным особенностям указанных постановок задач (общей и частных) для сложных систем относятся нелинейность, параметрические флуктуации, принципиальная неизмеряемость части координат. В связи с этим, возникает более общая постановка задачи: по наблюдаемым данным восстановить как функцию системы $f(\cdot)$, так и ее входные неизмеряемые координаты (в частности, задача построения наблюдателя для «скрытых» переменных в сложных технических системах (см. п. 7.5)).

В силу невозможности аналитического разрешения задачи в постановке (1.2) дадим интерпретацию общей постановки задачи (1.1) с точки зрения теории распознавания образов.

Интерпретация общей постановки задачи. Обозначим через Ω_0 - множество реальных состояний реальной ДО, Ω_Y - множество реализаций состояний ДО, отражающих математическую (или иную) модель, $\Omega_Y \subseteq \Omega_0$, Ω_{TS} - подмножество дискретных наблюдений реализаций состояний ДО, $\Omega_{TS} \subseteq \Omega_Y$, $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_I\}$ - множество образов (классов) состояний ДО, $\Omega \subseteq \Omega_{TS}$.

Первое предположение. Существует отображение $f: \Omega_0 \rightarrow \Omega_Y$ - математическая модель состояний реальной ДО (дифференциальное уравнение, случайный процесс и т.д.).

Второе предположение. Существует (неизвестное, однозначное) отображение $g: \Omega_0 \rightarrow \Omega_{TS}$ - модель состояний ДО (реальные значения состояний реальной ДО отображаются в дискретные подмножества $\Omega_i, i = \overline{1, I}$).

По заданной обучающей выборке в виде множества значений пар объект-метка состояний ДО $\Theta_Y = \{y_j^l, p_j^l\}_{j=1}^N \subset \Omega_{TS}$, где $l = \overline{1, k}$ - число серий наблюдений (реализаций) ДО, требуется решить задачи:

1) построить алгоритм распознавания (АлР) $A(\Theta_Y): \Omega_{TS} \rightarrow \Omega$ (возможно, множество АлР), осуществляющему оценку метки состояний ДО $d \in I_\Omega = \{1, \dots, I\}$ по совокупности наблюдений (n -окну) $y(j, n) = (y_{j-n}, y_{j-n+1}, \dots, y_j)$, $n \geq 1, j \in \{n, \dots, T\}$;

2) построить вектор оценок принадлежности (апостериорных вероятностей) к каждому из состояний ДО $\{p(d/y(j, n)), d \in I_\Omega\}$;

3) синтезировать итоговое решающее правило, согласующее частные решения пп.1, 2, и минимизирующее критерий качества (точности) распознавания Q , под которым понимают средний риск потерь [41], или вероятность неправильного распознавания;

4) оценить обобщающую способность множества построенных АлР на основе выбранного критерия качества (скользящий контроль, информационные критерии, максимальное правдоподобие, принадлежность решения парето-

оптимальному множеству).

1.4. Пример информационной модели оценивания состояний сложного объекта на основе временных рядов

В данном разделе описывается уровень математических и эвристических методов, применяемых для распознавания состояний ДО, включающий методы теории распознавания образов, теории информации, теории статистических оценок и логических методов, теории многокритериальных решений.

Подход к анализу ДО на основе временных рядов, развиваемый в данной работе, состоит в реализации следующих задач: 1) задачи сегментации, или выделения условно «однородных» фрагментов на основе анализа доступных реализаций фазовых траекторий и априорной информации; 2) задачи признакового описания выделенных фрагментов фазовых траекторий (временных рядов), соответствующих состояниям ДО; 3) задачи распознавания состояний ДО.

Поскольку временной ряд представляет собой дискретную реализацию некоторого случайного процесса, то подход к разметке временных рядов связан с задачей обнаружения моментов изменения свойств случайных процессов (в частности, моментов смены одного участка квазистационарности на другой), или с задачей обнаружения одного или нескольких моментов изменения (разладок) случайного процесса.

Существующие методы (и реализующие их алгоритмы) сегментации ВР удобно разбить на три основных группы.

1. Эвристические методы, основанные на предположении характерных свойств значений признаков состояний, удовлетворяющих некоторому заданному критерию однородности (функции структурной сложности) значений внутри фрагмента фазовых траекторий как состояния ДО. К эвристическим методам относятся алгоритмы сегментации по заданному критерию, алгоритмы кусочной и частичной аппроксимации с помощью многочленов и по некоторой системе базисных функций.

2. Методы на основе математической модели случайного процесса, соотношенного с поведением ДО и описывающего квазистационарные фрагменты фазовых траекторий с помощью семейства функций плотности распределения вероятности, а именно каждый квазистационарный участок кривой описывается вектором параметров, который и является признаком формы данного фрагмента [29, 30, 127, 128].

3. Методы на основе математической модели случайного процесса, описывающего и особенности однородных фрагментов и модель механизма чередования состояний (обычно в виде некоторого потока случайных событий). Известен подход к структурной обработке сигналов, использующий в качестве модели механизма чередования случайных событий конечную марковскую цепь пере-

ключений [120]. К основным недостаткам такой модели относится отсутствие возможности описывать фрагменты повторяющейся формы, т.к. если длительность фрагментов-состояний распределена на интервале $[1, \dots, \tau]$ и определяется диагональным элементом q_{ii} марковской матрицы Q , то это дает единственно возможное распределение вида $P(\tau_i=L)=(q_{ii})^{L-1} (1-q_{ii})$, $L=1, \dots, \tau$, $0 < q_{ii} < 1$, где τ_i - длина участка.

Определение 1.17. Информационная модель динамического объекта - это модель, созданная на языке блок-схемы с использованием методов, формул и графических интерпретаций, описывающая существенные для поставленной цели методы и алгоритмы анализа параметров и переменных объекта, связи между ними, оценивание неизмеряемых входов и выходов объекта и позволяющая путём подачи на модель информации об изменениях входных величин моделировать возможные состояния объекта.

Согласно [153] можно выделить несколько типов информационных моделей с точки зрения целевых установок: 1) для моделирования отклика объекта на внешнее воздействие; 2) для классификации внутренних состояний объекта; 3) для прогноза динамики изменения объекта; 4) для оценки полноты описания объекта; 5) для адаптивного управления объектом; 6) для оптимизации параметров объекта по отношению к заданной целевой функции.

Выбор методов и реализующих их алгоритмов оценивания состояний ДО на основе временных рядов в информационной модели (рис. 1.1) осуществлялся с целью решения задач распознавания состояний ДО, хотя представленные ниже методы и алгоритмы имеют самостоятельную ценность и могут применяться для решения других задач. Далее следует краткое описание отдельных блоков по указанным на рис. 1.1 (структурной схеме) номерам.

Сложный динамический объект (блок 1), как выше было обозначено, не имеет полного аналитического описания и полностью описывающей его модели. Совокупность траекторий ДО, получаемых при реализации различных начальных условий, образует обучающую выборку.

Под априорной информацией (блок 2) понимается информация, доступная до момента начала анализа ВР (запуска выбранного алгоритма распознавания).

При построении решающих правил и анализе их качества априорные сведения характеризуют: 1) пространство выбора; 2) природу случайных факторов; 3) оператор, определяющий характер взаимодействия между элементами пространства выбора и случайными факторами; 4) природу ВР; 5) требования к качеству решений.

Традиционная схема моделирования по стационарному временному ряду (блоки 3) достаточно хорошо изучена [33] и здесь обсуждаться не будет. На практике большинство процессов нестационарны и нелинейны, особенно в биологии, геофизике или экономике [30, 127]. В основе же большинства аналитиче-

ски обоснованных методов анализа и прогнозирования ВР лежит предположение либо о стационарности исследуемого СП, либо о «простоте» (малое число экстремумов по сравнению с числом точек) подбираемых простым перебором модельных трендовых функций.

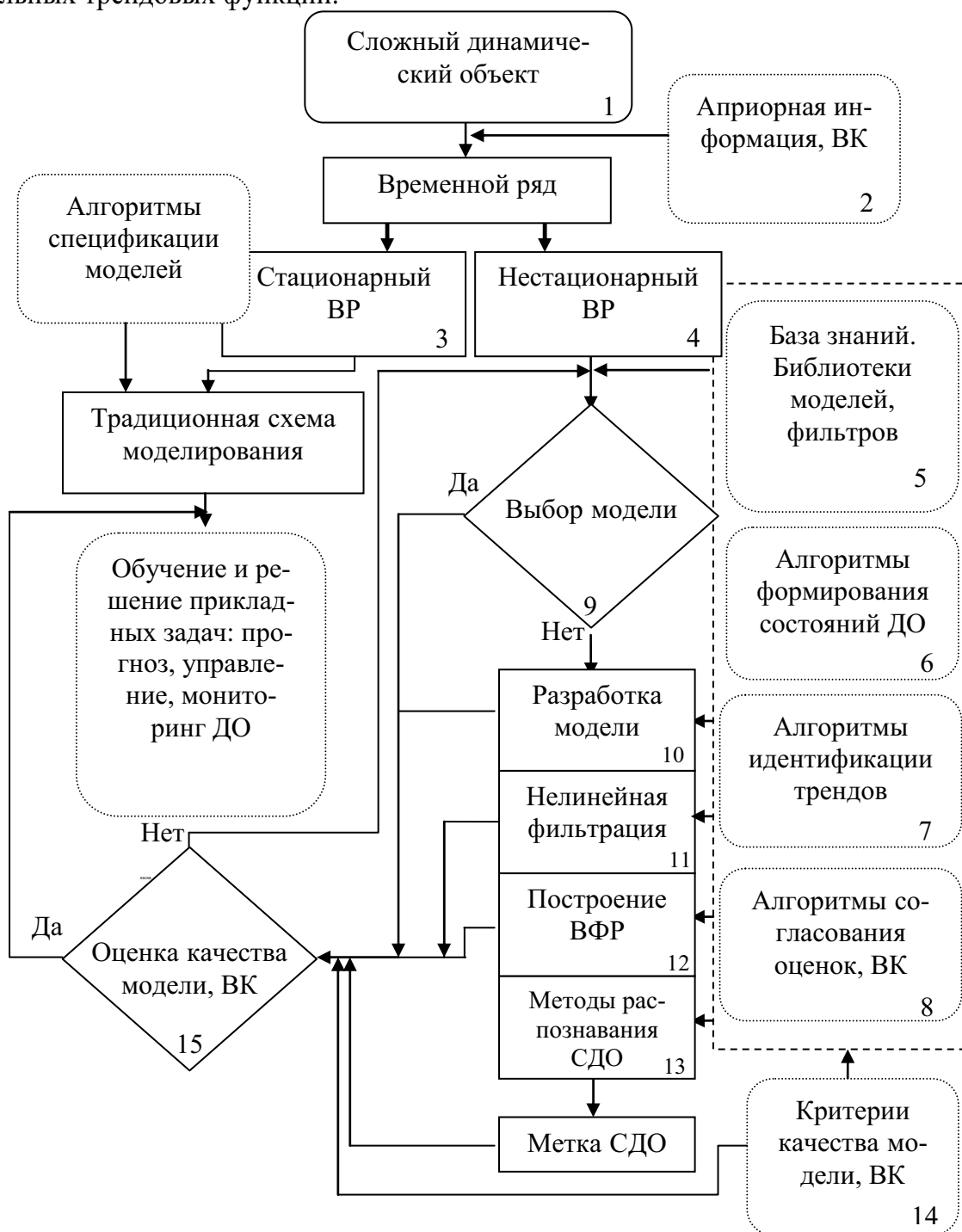


Рис. 1.1. Информационная модель оценивания состояний динамического объекта на основе временных рядов (ВР – временной ряд; VK – весовые коэффициенты; ДО – динамический объект; СДО – состояние динамического объекта; ВФР – выборочная функция распределения).

Так, классическая теория оптимальной фильтрации в настоящее время позволяет построить фильтры только для достаточного узкого класса шумов, близких к белому, поскольку общего формализованного подхода к идентификации шумов с другими законами распределения не существует. Отметим, что задача моделирования нестационарных ВР с нелинейным трендом (блок 4) на данный временной момент в общем виде в теоретическом плане не разрешена.

В информационной интеллектуальной модели (рис. 1.1) решение задачи распознавания состояний ДО на основе анализа сопровождающих его поведение ВР (с возможно нелинейными трендами), представлено блоками 5-14.

Расширяемые библиотеки моделей, фильтров (блок 5) содержат модели функциональных зависимостей (прецедентов при решении задач анализа ВР) и коэффициенты разностной схемы, им соответствующих, полученных на основе Z-преобразования [85, 157], а также могут быть пополнены программными модулями с фильтрами на основе базисных (ядерных [43]) функций.

Алгоритмы формирования состояний ДО (блок 6) связаны с задачей автоматической и экспертной разметки временного ряда [60, 139], исходя из особенностей ДО, выявленных по обучающей выборке (примеры алгоритмов разметки ВР приведены в пп. 2.4, 9.3).

Алгоритмы идентификации трендов (блок 7) реализуют решение задачи распознавания состояний ДО посредством формирования признаков формы состояний ДО, на основе которых возможно решение задачи реконструкции фрагмента временного ряда, содержательно соотношенного с состоянием динамического объекта (см. ниже п. 2.4.1). Алгоритмы базируются на аппроксимационных (сглаживающих) свойствах определенным способом формируемых обобщенных эталонов [85], сохраняющих геометрические особенности нелинейного тренда исследуемого временного ряда, на совмещении одновременного использования метода прямой экстраполяции [28], разностных схем и авторегрессионных моделей [25, 66], что позволяет учесть достоинства обоих методов.

Отметим, что алгоритм порождения моделей [150] служит для решения задачи восстановления нелинейной (не зашумленной) регрессии, этой же цели подчинен блок 11, реализующий алгоритмы восстановления нелинейной зашумленной регрессии на основе непараметрического подхода и на основе обобщенных эталонов [43, 85].

Алгоритмы согласования оценок (блок 8) опираются на байесовские оценки (гл. 5) и работы [75-79, 164]. Задача согласованного выбора информативных признаков состояний ДО ставится следующим образом. Пусть заданы L наборов признаков, выбранных на основе каких-либо методов (в том числе экспертных), показатели качества которых, удовлетворяют допустимым ограничениям. Требуется построить процедуру выбора оптимального набора признаков, относительно которого неопределенность выбора была бы минимальной.

Блоки 9, 14, 15 реализуются заданными критериями и процедурой непротиворечивого выбора, если критериев несколько. Так, если к частным критериям оценки выбираемой модели исследования отнести качество распознавания (f_1) и время распознавания (f_2) (методы гл. 2), полученных на контрольной выборке, то модификация метода МАИ (гл. 3) может служить основой оптимального по Парето выбора «наилучшей» пары (метод формирования признаков - алгоритм распознавания). Известно¹¹, что надежным критерием выбора алгоритма решения слабо формализованных задач является скользящий контроль.

Задача выбора модели анализа ВР тесно связана с проблемой классификации ВР на группы (классы, образы) схожих между собой ВР на основе экспериментальных данных. Классическим подходом к классификации рядов (в условиях отсутствия другой априорной информации, таких как характер шума, его свойства, наличие и вид тренда, наличие сезонных и циклических неслучайных компонент и пр.) является расчет попарных расстояний между всеми ВР и разбиение их на группы (кластеры) с помощью одного из известных алгоритмов кластерного анализа на основании матрицы расстояний (гл. 2, п. 2.4).

Блоки 10-13 (разработка модели) предполагает разработку новой комбинации алгоритмов (предобработки первичных измерений, формирования характеристических и классификационных признаков, распознавания состояний ДО, принятия итогового решения) на основе существующих алгоритмов и критериев качества, а также разработку новых алгоритмов. Методы, являющиеся модификациями известных методов распознавания образов на основе выявления новых закономерностей, изложены в гл. 2, 5.

Выводы. Сделан краткий обзор некоторых современных методов, используемых для анализа и прогнозирования временных рядов, и отмечены границы их применимости.

Представлена схема информационной модели для распознавания состояний сложного динамического объекта (рис. 1.1), описываемого нестационарным многомерным временным рядом с нелинейным трендом, с целью осуществления мониторинга и управления сложным объектом. Модель опирается на известные базовые и новые подходы к выявлению закономерностей в исходных данных при распознавании состояний сложного объекта. Приведены примеры постановок прикладных задач. Методы и алгоритмы, реализующие обсуждаемую информационную модель и изложенные ниже (гл. 2-9), апробированы на основе программного комплекса IReDDS [83, 84, 146].

¹¹ Например, работа Воронцова К.В. Обзор современных исследований по проблеме качества обучения алгоритмов // Таврический вестник информатики и математики. – 2004. – № 1. – С. 5–24.

«...весьма вероятно, что в очень многих случаях разумно изучение реальных явлений вести, избегая промежуточный этап их стилизации в духе представлений математики бесконечного и непрерывного, переходя прямо к дискретным моделям. Особенно это относится к изучению сложно организованных систем, способных перерабатывать информацию» /А.Н. Колмогоров/

Глава 2. ЭТАЛОННАЯ МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ И АЛГОРИТМЫ РАСПОЗНАВАНИЯ СОСТОЯНИЙ СЛОЖНОГО ДИНАМИЧЕСКОГО ОБЪЕКТА

В современных системах принятия решений крайне важной является проблема построения обобщенных моделей на основе обработки больших массивов экспериментальных данных, содержащих скрытые закономерности, выявление которых напрямую связано с успешным решением задачи искусственного интеллекта - обобщения информации.

Избавление от шума во временных рядах, по сути, приводит к искажению исходного признакового пространства, если исходная причина внешней «случайности» наблюдаемых данных временного ряда не стохастическая их природа, а фундаментальное следствие нелинейности закона обратной связи $Y_t = f(Y_{t-1}, \dots, Y_{t-p})$, в соответствии с которой генерируется временной ряд.

Так, генерация ВР по простейшей модели Фейгенбаума $Y_t = \alpha \cdot Y_{t-1} \cdot (1 - Y_{t-1})$ при $\alpha=4$ приведет к картине, внешне похожей на последовательность независимых реализаций некоторой случайной величины (белый шум).

В пользу применения теории распознавания образов и методов нелинейной динамики для исследования нестационарных ВР с нелинейным трендом высказывается и ряд отечественных авторов публикаций, посвященных методам анализа и прогнозирования временных рядов (см., например, обзоры в [29, 30, 158] и др.), и ряд зарубежных (напр., обзор в [168]). Неспособность отделить на базе наблюдаемых данных линейные модели со стохастической компонентой от нелинейных подрывает доверие к классическим моделям [12].

2.1. Общая схема оценивания состояний ДО по временным рядам методами распознавания образов и нелинейной динамики

Принципиальная сложность решения задачи распознавания состояний ДО методами распознавания образов вызвана следующими факторами:

1) необходимостью анализа обучающей выборки большого размера (для технических систем порядка более 10^4 – 10^6 значений временного ряда, а, как известно, задача поиска информативных наборов признаков является *NP*-полной);

2) обучением алгоритма распознавания и построением решающего правила на одной реализации случайного процесса, а принятием решения – на другой (возможно с другим уровнем и типом шумов);

3) реальностью времени, оцениваемой малыми порядками (~10 мкс для электронно-механических систем (ЭМС, п. 9.1));

4) возможной нелинейностью и нестационарностью зашумленного СП (1.1), причем причиной «случайного» поведения ДО может быть нелинейность;

5) возможным пересечением граничных состояний ДО.

Традиционное применение методов распознавания образов требует на основе имеющейся обучающей выборки (возможно состоящей из одной реализации СП) и дополнительной априорной информации относительно состояний ДО решить следующие подзадачи (рис. 2.1): 1) определить обобщенные характеристические (диагностические) признаки образов (классов) состояний ДО (построить информационное признаковое пространство (ИПП)); 2) на основе выделенного ИПП разработать и обучить алгоритмы распознавания состояний ДО; 3) оценить по контрольной выборке качество (надежность) всех АЛР состояний ДО; 4) построить (возможно, коллективную) процедуру для принятия итогового решения об отнесении наблюдаемых характеристик к одному из выделенных состояний ДО. Основу решающих правил составляют признаковые элементы, выбираемые для каждого состояния ДО из объектов обучающей выборки.

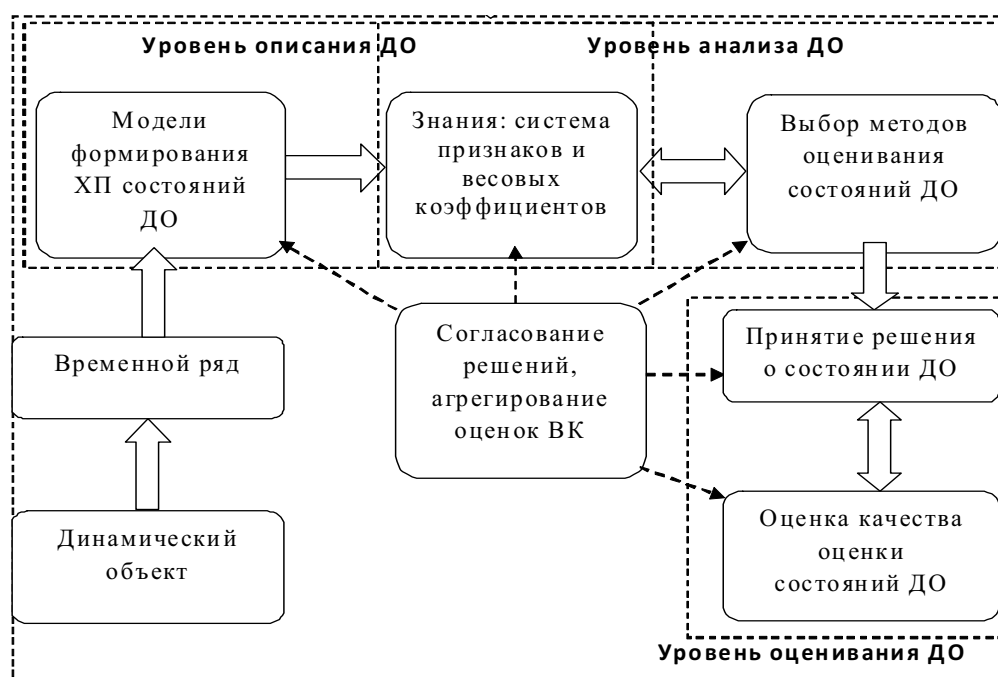


Рис. 2.1. Схема оценивания состояний ДО по временным рядам методами распознавания образов: ВК – весовые коэффициенты, ХП – характеристические признаки; широкие стрелки – поток данных и информации, пунктирные - поток информации

На рис. 2.1. представлена схема модели распознавания образов для решения задачи распознавания состояний ДО по временным рядам.

Основным математическим аппаратом нелинейной динамики (рис. 2.2) являются, как правило, нелинейные уравнения в частных (в обычных) производных. Такие уравнения могут отражать большой спектр различных режимов (состояний ДО), в отличие от линейных обыкновенных дифференциальных уравнений (ОДУ). Однако часть из этих уравнений не имеет аналитического решения, которое является наиболее ценным для анализа, поэтому используются многочисленные числовые алгоритмы решения таких уравнений и построения фазовых диаграмм. Поэтому совмещение методов распознавания образов, нелинейной динамики и математической статистики для исследования сложных ДО представляется перспективным [114, 115].



Рис. 2.2. Общая схема моделирования по временным рядам и место методов нелинейной динамики. В выделенных рамках – место настоящих исследований в общей схеме: оценивание и реконструкция состояний ДО

Реконструкция координат слабоформализованных ДО обеспечит решение следующих практически важных задач, список которых легко увеличить:

1) мониторинг, диагностика с целью решения задач прогнозирования и управления сложным объектом в выше определенном смысле;

2) определение наличия устойчивого многообразия (аттрактора) и его математической модели, необходимых для управления нелинейными с хаотической динамикой (например, по методу АКАР [70, 149] и I&I [2] (см. гл. 7);

3) классификация слабоформализованных многомерных динамических объектов (в т.ч. с присутствием немоделируемой динамики) и их распознавание с неполной информацией по выделенным признакам (зашумленные базы данных);

4) наблюдение (скрытых) величин в силу невозможности непосредственного их измерения по разным причинам: специфики модели и/или аппаратуры, возможности повреждения объекта (разрушающий контроль), неприемлемой дороговизны и пр. [3, 90];

5) осуществление верификации и адекватности моделей и их сравнение.

В данном разделе решена задача получения информативных признаков состояний ДО модели динамического объекта по его практической временной реализации (временному ряду) на основе обобщенных эталонов и принципа относительности расстояний [46, 58, 85] с целью применения классической схемы теории распознавания образов к решению задачи распознавания состояний ДО в реальном времени.

2.2. Метод обобщенных эталонов и решение задачи распознавания состояний динамического объекта

Применительно к ВР каждый из этапов метода распознавания образов имеет свои особенности. Однако вес неудачи на 1-м этапе (построение информационного признакового пространства) чрезвычайно велик, т.к. недостаточная информативность признаков не может гарантировать высокого качества классификации объектов из генеральной совокупности даже коллективом правил, или, другими словами, любой «хороший» алгоритм распознавания (состояний) покажет низкое качество, если исходная выборка нерепрезентативна. На проблему разумного соотношения числа признаков и числа объектов указывал А.Н. Колмогоров, а именно: если количество признаков превышает количество объектов, то велика вероятность обнаружения случайных псевдоинформативных подсистем [101].

Целью выявления закономерностей является производство нового знания для улучшения результатов принятия решений. К основным видам моделей, используемых для вывода и анализа нового знания, относятся модели классификации, регрессии, кластеризации, анализа ассоциаций, прогнозирования времен-

ных рядов, агрегирования (обобщения), обнаружение отклонений. Прежде чем изложить метод построения обобщенных эталонов, отметим следующие особенности традиционного применения методов кластеризации для построения ИПП (разбивающих выборку признаков на кластеры, состоящие из схожих признаков, и выделяющих в каждой группе по одному наиболее типичному представителю на основе введенной на множестве признаков метрики):

1) если в исходном множестве признаков будут присутствовать неинформативные признаки, то могут появиться кластеры, целиком состоящие из неинформативных признаков;

2) метрики используют информацию о попарном сходстве между признаками, но не решают проблему мультиколлинеарности признаков (набор попарно некоррелированных признаков может оказаться линейно зависимым);

3) классификация по принципу минимального расстояния пригодна, если значения признаков, описывающих объекты из разных образов, достаточно далеки друг от друга [55, 156].

Отметим также статистический кластер-алгоритм построения кластеров в предположении о нормальности распределений объектов в кластерах (что является естественным во многих практических приложениях), использующий кластер-критерий¹², предложенный А.М. Шурыгиным, и свойство оценок Л.Д. Мешалкина¹³ оценивать параметры распределения кластера, наибольшего по количеству точек. Заметим, что условия критерия К. Пирсона¹⁴ проверки гипотезы о совпадении центров двух многомерных нормальных совокупностей с известными ковариациями, различающихся сдвигом, являются непригодными для упомянутого кластер-анализа в силу жесткости используемых предположений.

Идея построения кластер-алгоритма, разбивающего выборку признаков на кластеры, состоящие из схожих признаков, заключается в следующем.

Кластер-критерием на уровне значимости α проверяется гипотеза о принадлежности выборки одному нормальному распределению. Если гипотеза не отвергается, то выборка с вероятностью $(1 - \alpha)$ содержит один кластер. Если гипотеза отвергается, то самый большой кластер вырезается эллипсоидом на уровне значимости α , и к оставшимся точкам применяется описанная процедура в цикле. Процедура выделения кластеров заканчивается, когда все точки рас-

¹² Шурыгин А.М. Статистический кластер-алгоритм // Математические методы распознавания образов: Сборник докладов. – М.: МАКС Пресс, 2007. – С. 241 – 242.

¹³ Meshalkin L.D. Some mathematical methods for the study of noncommunicable diseases // Proc. 6-th Intern. Meet. of Uses of Epidemiol. in Planning Health Services. – Yugoslavia, Primosten. – 1971. – Vol. 1. – Pp. 250–256.

¹⁴ Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling // Phil. Mag. – 1900. – Vol. 50. – Pp. 157–175.

пределены по кластерам, либо их количество невелико (α -остатки от проверки гипотез).

В отличие от большинства алгоритмов кластерного анализа данный статистический кластер-алгоритм не требует предварительного указания числа классов, на которое надо поделить совокупность точек наблюдения, и решает задачи со сложными наложениями и пересечениями нескольких классов (случай, когда признаки могут принимать одинаковые значения для объектов из разных образов).

1. Построение информационного признакового пространства. В качестве диагностических признаков для i -го состояния Ω_i , $i = \overline{1, I}$, принимается набор эталонов состояний ДО, построение которых осуществляется по одной из следующих схем: 1) выбор по обучающей выборке диагностических фрагментов соответствующего ВР (сегментов) и интервалов между ними [82]; 2) построение обобщенных g -эталонов (generalized etalon), суть которого – в замене исходного множества относительно близко расположенных (в определенном ниже смысле) в признаковом пространстве объектов одним g -эталонем. Остановимся на 2-й схеме, впервые примененной для ВР в [85] и оказавшейся весьма эффективной для нестационарных ВР с нелинейным трендом.

Постановка задачи преобразования исходной выборки в репрезентативную. Пусть дано некоторое множество элементов (объектов, эталонов) Ω_Y , представленное в виде объединения классов $\Omega_Y = \bigcup_{i=1}^I \Omega_i$. Каждый объект $Y_j \in \Omega_Y$ описывается системой признаков (координат, значений измерителей, датчиков), $Y_j = (a_{j1}, \dots, a_{jg})$. Имеется конечный набор объектов $U = \{Y_1, Y_2, \dots, Y_n\}$, $Y \subset \Omega_Y$, о каждом из которых известно, к какому классу он принадлежит. Необходимо построить новую систему обобщенных эталонов (g -эталонов) $Y' = (Y'_1, \dots, Y'_n)$, удовлетворяющих условиям [46]:

- а) размер системы g -эталонов должен быть меньше размера исходной обучающей выборки ($n' < n$);
- б) каждый g -эталон должен относиться к тому же состояний ДО, что и исходные эталоны, из которых он сформирован;
- в) качество получаемого решающего правила не должно ухудшаться по сравнению с решающим правилом, построенным по исходной выборке;
- г) обобщающая способность g -эталонов должна быть удовлетворительной (не меньше требуемой).

Общая схема распознавания нового объекта, не входящего в обучающую выборку, на основе метода эталонов [156] достаточно проста и известна: формируются по известным прецедентам эталоны каждого класса (образа) со значе-

ниями признаков, равных среднему арифметическому значений признаков объектов, образующих класс. Новый объект относят к тому классу, расстояние (как правило, евклидово) до эталона которого минимально. Недостатки данного метода также хорошо известны [58], в частности, метод имеет низкое качество распознавания в случае разной структуры классов (разного среднего внутриклассового расстояния).

Целесообразность применения FRiS-функции (как относительной меры сходства) при разной структуре классов показана в [13].

Определение 2.1. Функция F называется FRiS-функцией (Function of Rival Similarity), если для измерения степени сходства объекта C с конкурирующими объектами A и B в шкале отношений используются следующие соотношения: $F_{A/B} = (r_B - r_A) / (r_A + r_B)$ для сходства C с A в конкуренции с B , и $F_{B/A} = (r_A - r_B) / (r_A + r_B)$ для сходства C с B в конкуренции с A . Здесь r_A и r_B – расстояния от C до A и B , соответственно. При учете дисперсии внутриклассового расстояния $F_{1/2} = (r_2 d_1 - r_1 d_2) / (r_2 d_1 + r_1 d_2)$, $F_{2/1} = (r_1 d_2 - r_2 d_1) / (r_1 d_2 + r_2 d_1)$.

Применение FRiS-функции для данного случая заключается в следующем. Обозначим множества эталонов классов Ω_1 и Ω_2 через U_1 и U_2 , соответственно, $U_1 \cup U_2 = U$. Вычисляем значения степени относительного сходства исследуемого нового объекта y^* с ближайшими эталонами каждого из классов:

$$\gamma_{1/2} = F_{u/v} = \frac{r_v - r_u}{r_v + r_u}, \quad \gamma_{2/1} = F_{v/u} = \frac{r_u - r_v}{r_v + r_u},$$

где r_u и r_v – минимальные расстояния

от объекта y^* до ближайших u -го и v -го эталонов классов Ω_1 и Ω_2 , соответственно, с учетом их весовых коэффициентов. Решение принимается в пользу максимального значения степени сходства $i_0 = \arg \max(\gamma_{1/2}, \gamma_{2/1})$.

Методы отбора эталонов, используемых в метрических методах классификации (методы ближайших соседей, потенциальных функций, метод парзеновского окна), можно классифицировать (на настоящий момент) на три группы:

1) методы (Stolp, A-Stolp, FRiS-Stolp), основанные на FRiS-функции, суть применения которой означает получение оценок локальных плотностей классов в каждом объекте и вычисление отношений этих оценок; методы обладают хорошей обобщающей способностью;

2) метод опорных векторов SVM, в котором эталонами (опорными объектами) являются «пограничные» объекты; метод релевантных векторов RVM, отбирающий в качестве эталонов объекты, отстоящие от границы классов на определяемом расстоянии с целью избавления от шумовых объектов; в RVM методе число опорных объектов меньше по сравнению с SVM и обобщающая способность выше;

3) метод отбора эталонов для алгоритма ближайшего соседа, основанный на минимизации функционала полного скользящего контроля [47], недостатком которого является низкая эффективность по времени.

Нижеизложенный алгоритм 2.1, обладая достоинствами FRiS-функции, идейно восходит к методу потенциальных функций (вес эталонов зависит от расстояния до межклассовой границы), исследовался в работе [46] и обобщен для применения к временным рядам в совместной работе [85].

К особенностям методики формирования эталонов относятся следующие: 1) используется принцип относительности расстояний: объект с номером j исходной выборки образуют новый g -эталон определенного класса, если расстояние от объектов до выбранной точки ρ меньше расстояния от объектов до конкурирующей точки ρ^c ($r_{\rho j} < r_{\rho^c j}$); объект, расстояние от которого до начальной точки ρ больше расстояния между начальной ρ и конкурирующей ρ^c точками в новый g -эталон не включаются ($r_{\rho j} > r_{\rho\rho^c}$); 2) каждый класс может иметь несколько g -эталонов, выражающих его пространственную (геометрическую) индивидуальность в вычисляемых весовых коэффициентах, равных числу присоединенных элементов исходной выборки к данному эталону.

Без ограничения общности изложим суть алгоритма 2.1 для $\Omega = \{\Omega_1, \Omega_2\}$. Выбирается объект ρ (начальная точка формирования g -эталона), наиболее удаленный от всех объектов выборки; выбирается объект другого класса ρ^c (конкурирующая точка формирования g -эталона), наиболее близкий к точке ρ . Во множество U_ρ включаются объекты исходной выборки, расстояние до которых от точки ρ меньше расстояния до точки ρ^c (рис. 2.3), и не включаются объекты, расстояние от которых до начальной точки ρ больше расстояния между начальной ρ и конкурирующей ρ^c точками. Признаками g -эталона являются средние значения признаков вошедших в него объектов.

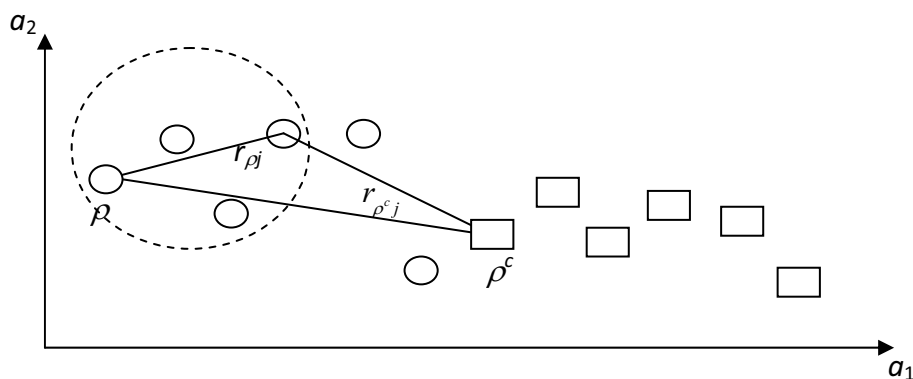


Рис. 2.3. Иллюстрация формирования g -эталона

Алгоритм 2.1. Формирование g -эталонов

Вход: Обучающая выборка $U = \{Y_1, Y_2, \dots, Y_n\}$, $Y_j = (a_{j1}, \dots, a_{jg})$ {значения признаков для каждого объекта Y_j }.

Выход: $U_G = \{G_1, G_2, \dots, G_m\}$, $m \leq n$, $G_j = \{b_{j1}, \dots, b_{jg}, w_j\}$, где a'_{kj} - усредненные значения признаков для каждого g -эталона Y_j , w_j - число объектов, вошедших в g -эталон G_j .

1: Для всех $i=1, \dots, n$, для всех $j=1, \dots, n$: найти расстояния:

$$r_{ij} = \sqrt{\sum_{e=1}^g (a_{ie} - a_{je})^2};$$

2: Пока выборка $U \neq \emptyset$ осуществлять действия:

3: найти начальную точку формирования g -эталона $\rho = \arg \max_{j=1, |\Omega_1|}^{|\Omega_2|} \sum_{j=1} r_{ij}$;

4: найти точку формирования конкурирующего g -эталона $\rho^c = \arg \min_{j=1, |\Omega_2|} r_{\rho j}$;

5: сформировать множество объектов $U_\rho = \{Y_j \mid r_{\rho j} < r_{\rho^c j}, r_{\rho j} < r_{\rho^c \rho}\}$;

6: зафиксировать новый g -эталон G_j на основе U_ρ со значениями признаков, равными $b_{j\rho} = w_\rho^{-1} \sum_{l=1}^{w_\rho} a_{jl}$, $w_\rho = |U_\rho|$;

7: удалить из U объекты U_ρ : $U := U \setminus U_\rho$.

Вернуть $U_G = \{G_1, G_2, \dots, G_m\}$, $G_j = \{b_{j1}, \dots, b_{jg}, w_j\}$.

Комментарий к алгоритму 2.1. В алгоритме 2.1 выбирается объект ρ («центр» g -эталона), наиболее удаленный $\left(\rho = \arg \max_{j=1, |\Omega_1|}^{|\Omega_2|} \sum_{j=1} r_{ij} \right)$ от всех объектов выборки $\Omega = \{\Omega_1, \Omega_2\}$; выбирается объект другого класса $\rho^c = \arg \min_{j=1, |\Omega_2|} r_{\rho j}$, наиболее близкий к точке ρ (конкурирующая точка формирования g -эталона). Во множество U_ρ включаются объекты по правилу: 1) расстояние от точки ρ до объектов меньше расстояния до точки ρ^c ($r_{\rho j} < r_{\rho^c j}$); 2) расстояние от точки ρ до объектов меньше расстояния между начальной ρ и конкурирующей ρ^c точками. Признаками g -эталона являются усредненные значения признаков объектов, его образующих.

На рис. 2.4 веса эталонов интерпретированы как площади фигур-объектов. Эталон с небольшим весом могут быть удалены из сформированного множества как нетипичные (или зашумленные) по пороговому принципу. Так, на рис. 2.4. *d*) - обобщенные эталоны после пороговой фильтрации (удалены эталоны с весом, меньше 17).

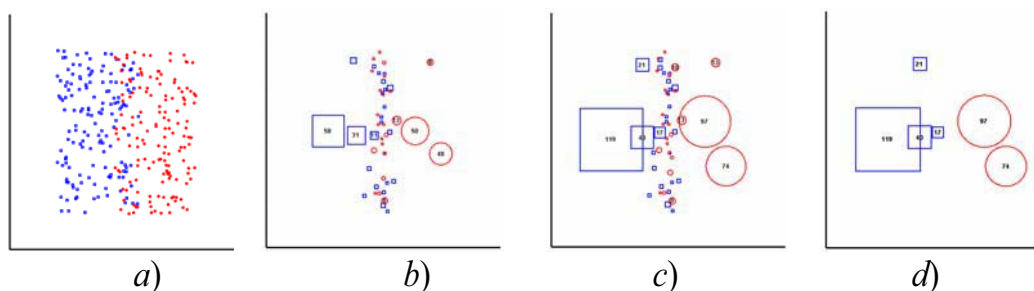


Рис. 2.4. а) 150 объектов в каждом классе с 30% пересечением; б) обобщенные эталоны: 24 эталона в 1-м и во 2-м классах; в) 26, 25 эталонов в 1-м и во 2-м классах при добавлении по 100 объектов; д) эталоны после фильтрации

В таблице 2.1 приведены результаты экспериментальных исследований зависимости эффективности решающих правил от площади пересечения состояний ДО для обучающей выборки размером 10000 объектов. Результаты являются средними по 50-ти экспериментам. Для построения решающих правил использовался алгоритм 2.2.

Анализ результатов исследований показал, что g-эталон с малым весом можно считать шумом и удалять, тем самым приведенный алгоритм можно использовать и как фильтр исходной выборки и как способ формирования взвешенных обобщенных признаков, выявляющих типичные особенности первичных измерений показателей состояний.

Таблица 2.1. Зависимость эффективности решающих правил от площади пересечения состояний при 30%-м шуме

Площадь пересечения состояний	Эффективность распознавания состояний с применением эталонов		Время распознавания по g-эталонам /по обычным, мкс
	g-эталон	Эталон обычного усреднения	
0	99.9%	99.9%	
10	99.8%	99.3%	0.26/0.87
20	98.1%	97.0%	0.53/4.36
30	97.4%	86.4%	0.75/6.97
40	96.5%	54.1%	1.11/22.58
50	95.2%	34.7%	1.51/27.77

2. Оценивание информативности признаков. При вводе термина «информативные признаки» подразумевается наличие некоторой информационной характеристики, описывающей совокупность признаков (информационный критерий Акаике, энтропийный критерий Шеннона). Предполагается, что на мно-

жестве измеренных характеристик существуют подмножества признаков наиболее полно выражающих суть некоторой заданной задаче анализа данных.

К детерминированным показателям информативности признаков, используемых в ИМ, относятся:

1. Традиционно в качестве критерия информативности признаков используется доля правильно распознанных объектов обучающей выборки в процедуре Cross – Validation (см. определение в п.1.2).

2. Показатель относительного сходства $\sum_{j=1}^n (r_j^2 - r_j^1) / (r_{\max} - r_{\min})$, где n – объем обучающей выборки, r_j^1 и r_j^2 – расстояния j -го объекта из обучающей выборки до ближайшего объекта своего образа и до ближайшего объекта образа-конкурента, соответственно; r_{\max} и r_{\min} – максимальное и минимальное расстояния между объектами обучающей выборки [9].

3. Критерий, использующий FRiS-функцию [13] (определение 2.1).

4. Показатели эффективности (являющиеся числовым выражением информативности признаков) на основе нелинейной скаляризации (гл. 3) нескольких критериев.

5. Меры относительной важности признаков, основанные на мультимножествах (пп. 4.2, 4.7, формулы (4.11))

6. Метрические и неметрические методы многомерного шкалирования [25]. В данной работе приведена модификация (ММАИ) метода анализа иерархий (МАИ) и доказаны ее свойства (гл. 3), которая позволяет корректно ранжировать признаки в динамически изменяющихся наборах признаков.

Статистическими показателями информативности признаков, используемых в ИМ, являются следующие.

1. Для случая нормального распределения значений признаков применяется критерий Фишера в виде отношения $Q = |m_1 - m_2| / (d_1 + d_2)$, где m_1 , m_2 и d_1 , d_2 – математические ожидания и дисперсии первого и второго образов, соответственно.

2. Мера близости (метрика) на основе коэффициента линейной корреляции $r(z_1, z_2)$: $\rho(z_1, z_2) = 1 - |r(z_1, z_2)|$, $r(z_1, z_2) = \sum_{i=1}^n \bar{z}_1(y_i) \cdot \bar{z}_2(y_i)$, где $\bar{z}_1(y_i)$, $\bar{z}_2(y_i)$ – нормированные и центрированные значения признаков z_1 , z_2 для объектов выборки $y = (y_1, \dots, y_n)$. Расстояние $\rho(z_1, z_2) = 0$ тогда и только тогда, когда признаки связаны линейной зависимостью.

2. Меры важности признаков на основе энтропийного критерия Шеннона и дивергенции Кульбака (гл. 5).

3. Алгоритмы распознавания состояний ДО. Классические детерминированные (метрические) методы РО (метод эталонов, метод «ближайших соседей») модифицированы и адаптированы к специфике обучающей выборки рассматриваемых задач. Изложим используемые в информационной модели (рис. 0.1, 1.1) их модификации на основе обобщенных эталонов.

Алгоритм распознавания состояний ДО на основе метода эталонов и функции относительного сходства (FRiS-функции). Основу решающего правила составляют эталонные элементы, выбираемые для каждого образа из объектов обучающей выборки по алгоритму выбора эталонов-подпоследовательностей, или g-эталонов. Для распознавания нового объекта (значения сигнала) определяются расстояния от него до всех эталонов, выбираются ближайшие два эталона из разных классов, и по расстояниям до них вычисляются значения FRiS-функции. Решение принимается в пользу того класса, значения FRiS-функции для которого максимально.

Алгоритм 2.2. Принятие решения по g-эталонам и функции относительного расстояния

Параметры оптимизации: весовые коэффициенты $\{w_l^i, l = \overline{1, N}, i = \overline{1, I}\}$; вес класса.

Вход: Значение сигнала y ; Набор эталонов $B^{(i)} = \{B_j^{(i)}, j = \overline{1, n_i}\}$, $B_j^{(i)} = (b_{j_1}^{(i)}, \dots, b_{j_N}^{(i)})$, для каждого i -го распознаваемого состояния $\Omega_i, i = \overline{1, I}$; n_i – число эталонов для i -го состояния ДО; эталонное внутриклассовое расстояние δ_i .

Выход: Номер i_0 состояния ДО, в пользу которого принимается решение.

1: вычислить расстояния $\{d(y, B_j^{(i)}), j = \overline{1, n_i}, i = \overline{1, I}\}$ от (объекта) y до всех имеющихся эталонов $B^{(i)} = \{B_j^{(i)}, j = \overline{1, n_i}\}$ по формуле

$$d(y, B_j^{(i)}) = \sqrt{\sum_{l=1}^N (w_l^i)^2 (y_l - b_{j_l}^{(i)})^2};$$

2: найти ближайшие два эталона из разных образов $B^{(i_1)}, B^{(i_2)}$: $i_1 = \arg \min_{i=1, I; j=1, n_i} d(y, B_j^{(i)}), i_2 = \arg \min_{i=1, I; i \neq i_1; j=1, n_i} d(y, B_j^{(i)});$

3: формировать FRiS-функции по формулам: $F_{i_1/i_2} = (r_{i_2} - r_{i_1}) / (r_{i_1} + r_{i_2})$ для сходства y с классом S_{i_1} в конкуренции с S_{i_2} , и $F_{i_2/i_1} = (r_{i_1} - r_{i_2}) / (r_{i_1} + r_{i_2})$ для сходства y с S_{i_2} в конкуренции с S_{i_1} . Здесь r_i и r_j – расстояния от y до S_i и S_j , соответственно.

Вернуть: $i_0 = \arg \max(F_{i_1/i_2}, F_{i_2/i_1})$.

На основе алгоритма агрегирования полученных «частных» (возможно, и противоречивых) голосов за номер (метку) состояния, к которому предлагается отнести распознаваемый сигнал $y(t)$ по отдельному АлР, осуществляется синтез итогового решения в пользу одного состояний ДО. Алгоритмы агрегирования совокупности «голосов» за метку состояния ДО будут приведены далее и основаны на ММАИ [76-79] и согласовании байесовских оценок [164].

Пример 2.1. Покажем работоспособность алгоритмов для обработки первичной информации и для итогового распознавания состояний динамического объекта на решении прикладной задачи при распознавании состояний асинхронного двигателя (подробно гл.4, 9).

В табл. 2.2, 2.3 приведены результаты распознавания состояний ДО по следующим алгоритмам: АлР 1 на основе метода обычных эталонов [156] и FRiS-функции; АлР 2 на основе традиционного метода эталонов и ближайших соседей; АлР 3 на основе метода построения g -эталонных и FRiS-функции; АлР 4 на основе коллективного решающего правила, FRiS-функции и применении весовых коэффициентов АлР, полученных на контрольной выборке. Эффективность распознавания в процентах оценивались по распознаванию 10000 объектов. Через $U(-X, X)$ обозначен равномерный шум на интервале $(-X, X)$ с нулевым средним.

Таблица 2.2. Результаты распознавания для 4-х алгоритмов РО

Число эталонов в классе-состоянии				Тип и уровень шумов	Эффективность распознавания, %			
Номер состояния					АлР 1	АлР 2	АлР 3	АлР 4
1	2	3	4					
5	5	5	5	нет шума	95.51	93.89	94.77	95.51
5	5	5	5	$U(-320, 320)$	47.12	25.42	68.51	39.00
2	101	26	370	нет шума	98.67	99.45	99.67	99.67
2	101	26	370	$U(-320, 320)$	94.84	90.97	98.51	94.16

Таблица 2.3. Зависимость времени распознавания от числа эталонов

Число эталонов в классе-состоянии				Время распознавания, мкс			
1	2	3	4	АлР 1	АлР 2	АлР 3	АлР 4
1	1	1	1	0.16	0.26	0.87	1.42
5	5	5	5	0.24	0.53	4.36	5.21
10	10	10	10	0.51	0.75	6.97	8.08
20	20	20	20	0.57	1.11	22.58	24.34

30	30	30	30	1.00	1.51	27.77	30.27
4	184	47	672	4.53	6.56	1698.91	1738.28

Как следует из таблиц 2.2, 2.3 существует такой набор эталонов по состояниям (см. строки в табл. 2.2 с типом шумов $U(-320, 320)$), который дает удовлетворительное сочетание (качество распознавания-время распознавания) даже при большом уровне шума (дисперсии). Время распознавания в табл. 2.3 указано суммарное, включающее и время формирования эталонов.

4. Оценивание качества алгоритмов распознавания состояний ДО.

Традиционно в теории распознавания образов качество распознавания оценивается значением отношения числа правильных ответов к общему объему контрольной выборки на скользящем контроле. Как показано в ряде работ школы В.К. Воронцова, все информационные критерии уступают в точности оценивания качества алгоритмов распознавания методу скользящего контроля.

Алгоритмы распознавания состояний ДО на основе детерминированных и статистических методов, методы формирования информативного подпространства признаков порождают множество решающих правил, и вопрос о целесообразности применения того или иного решающего правила для каждой задачи распознавания состояний ДО должен решаться отдельно.

Отмечено, что совместное использование в едином решающем правиле разнотипных моделей является перспективным способом наиболее полного учёта априорной информации, заключающейся в обучающей выборке.

Корректный выбор оптимального по Парето метода (алгоритма 2.3) распознавания состояний ДО A_l по нескольким информационным критериям предлагается осуществлять по нижеизложенной нелинейной модификации (гл. 3) широко известного метода МАИ [141], являющегося основой многих систем поддержки принятия решения, ведущей из которых является зарубежный пакет ExpertChoice.

Алгоритм 2.3. Выбор Парето-оптимального алгоритма распознавания

Вход: показатели качества C_s и их назначенный весовой коэффициент c_s ,

$\sum_{s=1}^v c_s = 1$; множество оцениваемых алгоритмов распознавания R_j , $j = \overline{1, g}$, а так-

же значения λ_j^s показателей качества C_s для алгоритмов распознавания R_j .

Выход: вектор приоритетов алгоритмов распознавания (V_1, V_2, \dots, V_g) и номер «наилучшего» алгоритма распознавания.

1: Строим матрицу парных сравнений (МПС)

$\mathbf{A}^s = \|a_{ij}^s\|_{g \times g}$, $a_{ij}^s = \lambda_i^s (\lambda_j^s)^{-1}$ на каждом из v этапов (по числу v показателей

качества алгоритмов C_s). Результатом s -го этапа является вектор нормализованных значений ВКА (весовых коэффициентов альтернатив) (алгоритмов) – $W_s = (w_1^s, w_2^s, \dots, w_g^s)$.

2: Формируем всевозможные векторы $w_{ij}^s = (w_{ij}^s(i), w_{ij}^s(j))$ локальных ВКА уровня 1: $w_{ij}^s(i) = \frac{w_i^s}{w_i^s + w_j^s}$, $w_{ij}^s(j) = \frac{w_j^s}{w_i^s + w_j^s}$, $s = \overline{1, v}$, $j = \overline{1, g}$.

3: Формируем матрицу $\mathbf{W} = \|w_{ij}\| = \|(w_{ij}(i), w_{ij}(j))\|$, где векторы $(w_{ij}(i), w_{ij}(j))$ – локальные нормализованные ВКА z_i, z_j уровня 2 относительно всей совокупности мер относительной важности признаков, компоненты которых находим по формулам: $w_{ij}(i) = \sum_{s=1}^v c_s w_{ij}^s(i)$, $w_{ij}(j) = \sum_{s=1}^v c_s w_{ij}^s(j)$.

4: Формируем $u_i = \sum_{l=1}^g w_{il}(i)$, $V_i = u_i / \sum_{l=1}^g u_l$, $i = \overline{1, g}$.

Вернуть: вектор приоритетов (V_1, V_2, \dots, V_g) , номер «наилучшего» алгоритма, равного $\arg \max_{i=1, g} V_i$.

5. Принятие итогового решения. На основе предложенного алгоритма агрегирования полученных «частных» (возможно, и противоречивых) голосов за номер (метку) класса, к которому предлагается отнести распознаваемое значение (фрагмент) сигнала $y(t)$ по отдельному АлР, возможен синтез итогового решения по нескольким АлР. В данном случае роль критериев играют алгоритмы распознавания, роль альтернатив – распознаваемые состояния. Иллюстративный пример применения алгоритма 2.3. изложен в п.3.3.

В следующем разделе решена задача получения эволюционной математической модели динамического объекта по его временной реализации (временному ряду) на основе обобщенных эталонов и принципа относительности состояний [13, 46, 85], а также на основе свойств разностных и авторегрессионных схем.

2.3. Реконструкция формы состояний стохастического объекта на основе модели эталонов

При исследовании реальных сложных ДО в силу затруднительного (невозможного) непосредственного исследования его структуры анализ ДО осуществляется на основе какого-либо сигнала, сопровождающего ДО, и рассматриваемого как многомерный временной ряд. Анализ таких сигналов на основе тради-

ционных методов (Фурье-преобразований, Wavelet-разложений и т.д.) существенно осложняется нелинейностью динамических характеристик и возможной неопределенностью (динамической и/или измерительной).

Реализация процедуры реконструкции необходима для решения широкого круга самостоятельных задач, связанных с необходимостью получения моделей сложных нелинейных ДО. Напомним, что в нелинейной динамике наряду с термином «идентификация» применяется термин «реконструкция».

Классической основой реконструкции является задача аппроксимации точек на плоскости (x, y) функцией $y=f(x)$ [59]. В современных исследованиях стоит вопрос описания сложных (нелинейных, слабоформализованных) процессов, поэтому появляется необходимость построения по экспериментальным данным разностных уравнений $y_{t+1}=g(y_t)$.

Построенные по рядам эмпирические модели имеют большой интерес для организации прогноза дальнейшего поведения, для оценки адекватности модельных представлений о физической сущности процесса, и являются дополнением традиционных методов анализа ДО (спектральный, автокорреляционный, информационный и т.д.) для оценки структуры фазовых портретов, бифуркационных диаграмм и особенностей пространства параметров.

При этом традиционная основная идея остаётся неизменной и заключается в подгонке коэффициентов аппроксимирующих функций f_i по точкам в многомерном фазовом пространстве $\{y_i\}$, что по силам только вычислительной технике с современной производительностью и требует разработки специальных оптимизирующих алгоритмов, определения стартовых значений параметров оптимизации и их весовых коэффициентов.

Как известно, не существует общего формализованного подхода для «автоматического» обнаружения тренда во временном ряде и его моделирования (идентификации). Если временные ряды содержат значительную ошибку, то нередко первым шагом выделения тренда является сглаживание.

Сглаживание всегда включает некоторый способ локального усреднения данных, при котором несистематические компоненты взаимно погашают друг друга. К общим методам сглаживания относятся: метод скользящего среднего, в котором каждый член ряда заменяется простым или взвешенным средним k соседних членов, где k - ширина «окна» [33]; метод медианного сглаживания значений, попавших в «окно»; метод бикубических сплайнов; метод сглаживания методом наименьших квадратов, взвешенных относительно расстояния; отрицательного экспоненциально взвешенного сглаживания; метод подгонки функции (если известен вид функциональной зависимости).

Для приближения ВР линейной функцией при существовании явной монотонной нелинейной компоненты, данные вначале преобразуют для устранения нелинейности применением логарифмического, экспоненциального или полиномиального преобразования данных.

Каждый из методов имеет свои условия применения, и, соответственно, недостатки и достоинства. Так, к преимуществу медианного сглаживания относится большая устойчивость к выбросам (имеющимся внутри окна) по сравнению со скользящим средним, а к недостатку – получение более «зубчатых» кривых при отсутствии явных выбросов и отсутствие возможности использовать весовые коэффициенты. Ряды с относительно небольшим количеством наблюдений и систематическим расположением точек могут быть сглажены с помощью бикубических сплайнов. Методы МНК и экспоненциальное сглаживание применяют при больших ошибках измерения для фильтрации от шума и преобразования данных в относительно гладкую кривую. Более сложным и более эффективным является способ фильтрации, основанный на методе наименьших квадратов. В пределах окна экспериментальные данные аппроксимируют каким-либо полиномом (например, квадратичной или кубической функцией). Широко распространен метод Савицки и Голэя (Savitzky-Golay smoothing [19]), в котором задается определенное число экспериментальных точек k (ширина окна) и методом полиномиальной регрессии определяется сглаженное значение для каждой точки и окно передвигается дальше по всей кривой.

Рассмотрим подход к сглаживанию (фильтрации) и реконструкции тренда фрагментов временных рядов, соотнесенных с состояниями стохастических (сложных) ДО. Подход реализуется алгоритмами формирования обобщенных эталонов состояний многомерного временного ряда, сглаживания временного ряда на их основе, формирования признаков состояний как формы функциональной зависимости на основе совмещения методов прямой экстраполяции, разностных схем и обобщенных эталонов состояний ряда.

Пусть по обучающей выборке построено множество g -эталонов (алгоритм 2.1), усредненных по числу доступных реализаций.

Фильтрующий (сглаживающий) алгоритм 2.4 без ограничения общности излагается для 2-х состояний (классов). Общая схема формирования фильтрующих g -эталонов следующая. Для нового измерения сначала распознается номер состояния (см., например, алгоритм 2.2). Далее выполняется проверка на близость к g -эталонам «своего» и конкурирующего классов с учетом весовых коэффициентов g -эталонов. Если добавляемый объект находится ближе к g -эталону «своего» класса, то выполняется корректировка значений признаков этого g -эталона. Если добавляемый объект находится ближе к g -эталону «чужого» класса, то он объявляется новым g -эталонem с единичным весом. По окончании процесса «незначимые» g -эталоны с весовыми коэффициентами, меньшими порогового, удаляются.

Степень «близости» измеряется на основе введенного коэффициента относительного сходства между наблюдаемым объектом Y_i^* и g -эталонem

$G_j = (a_{j1}, \dots, a_{jg})$: $\lambda_j = w_j \left[\sum_{e=1}^g (a_{le}^* - a_{je})^2 \right]^{-1}$, обратно пропорционального расстоянию между объектом и g -эталонном (как в методе потенциальных функций).

Алгоритм 2.4. Сглаживание временного ряда

Вход: множество g -эталонных $U_G = \{G_1, G_2, \dots, G_m\}$, $G_j = \{a_{j1}, \dots, a_{jg}, w_j\}$, $j = \overline{1, m}$, и множество наблюдаемых объектов $Y_l^* = \{a_{l1}^*, \dots, a_{lg}^*\}$, $l = \overline{1, L}$, в виде измерений $(g-1)$ -мерного временного ряда (признак a_{lg}^* – значение времени измерения Y_l^*), а также – параметр фильтрации w_0 ;

Выход: множество сглаживающих пороговых g -эталонных $U'_G = \{G'_1, G'_2, \dots, G'_{m'}\}$, $G'_j = \{a_{j1}, \dots, a_{jg}, w_j\}$, $j = \overline{1, m'}$, $m' < m$.

1: Положить $G'_j := G_j$, $m' := m$, $j = \overline{1, m'}$, $U'_G := U_G$.

2: Вычислить коэффициенты относительного сходства между наблюдаемым объектом Y_l^* и всеми g -эталонами G'_j состояния Ω_i^* :

$$\lambda_j = w_j \left[\sum_{e=1}^g (a_{le}^* - a_{je})^2 \right]^{-1}.$$

3: Найти ближайшие g -эталонные «своего» класса $\alpha = \arg \max_{j=1, m, G'_j \subset \Omega_i^*} \lambda_j$, $i^* = \arg \max (F_{i_1/i_2}, F_{i_2/i_1})$ и класса-конкурента $\beta = \arg \max_{j=1, m, G'_j \not\subset \Omega_i^*} \lambda_j$. Эталон $G'_j \subset \Omega_i$, если $t_{i-1} \leq a_{jg} < t_i$, где интервал (t_{i-1}, t_i) соотнесен с состоянием Ω_i .

4: Принять решение $Y_l^* \in G'_\alpha$, и значения признаков g -эталона G'_α пересчитать: $a_{\alpha j} := \frac{a_{\alpha j} \cdot w_\alpha + a_{lj}^*}{w_\alpha + 1}$, $w_\alpha := w_\alpha + 1$, если $\lambda_{l\alpha} - \lambda_{l\beta} \geq \delta$, δ – параметр оптимизации.

5: Принять решение $Y_l^* \in G'_\beta$, и значения признаков g -эталона G'_β пересчитать: $a_{\beta j} := \frac{a_{\beta j} \cdot w_\beta + a_{lj}^*}{w_\beta + 1}$, $w_\beta := w_\beta + 1$, если $\lambda_{l\alpha} - \lambda_{l\beta} < -\delta$.

6: Принять решение $m' := m' + 1$, $G'_{m'} := Y_l^*$, $w'_{m'} := w'_l$, если $(|\lambda_{l\alpha} - \lambda_{l\beta}| < \delta)$.

7: Удалить из U'_G объекты G'_u : $U'_G := U'_G \setminus G'_u$, $w'_l < w_0$.

Вернуть: $U'_G = \{G'_1, G'_2, \dots, G'_{m'}\}$.

Иллюстрация работы алгоритма 2.4 приведена в конце раздела.

2.3.1. Метод моделирования состояний ДО на основе обобщенных эталонов и параметризации модели авторегрессии

В настоящее время для моделирования и прогнозирования популярны непараметрические методы в силу сложности (невозможности) аналитической параметризации априорно нелинейных по параметрам временных рядов. Однако, в непараметрических методах невысока точность прогнозирования, не строится теоретическая модель явления или процесса.

Параметрические методы демонстрируют «общую картину» развития процесса, дают, как правило, большую точность прогноза.

Алгоритм реконструкции фрагмента временного ряда, содержательно соотношенного с состоянием динамического объекта, основан на аппроксимационных (сглаживающих) свойствах g -эталонов, на одновременном применении метода прямой экстраполяции [28] и авторегрессионных моделей, что позволяет учесть достоинства обоих методов.

Суть метода прямой экстраполяции состоит в определении параметров для всех моделей-претендентов и выборе одной из них в смысле заданного критерия качества. Однако, как показано в [25], выбор наиболее значимой модели (в смысле F -критерия) из нескольких заданных с разным функциональным видом неоднозначен.

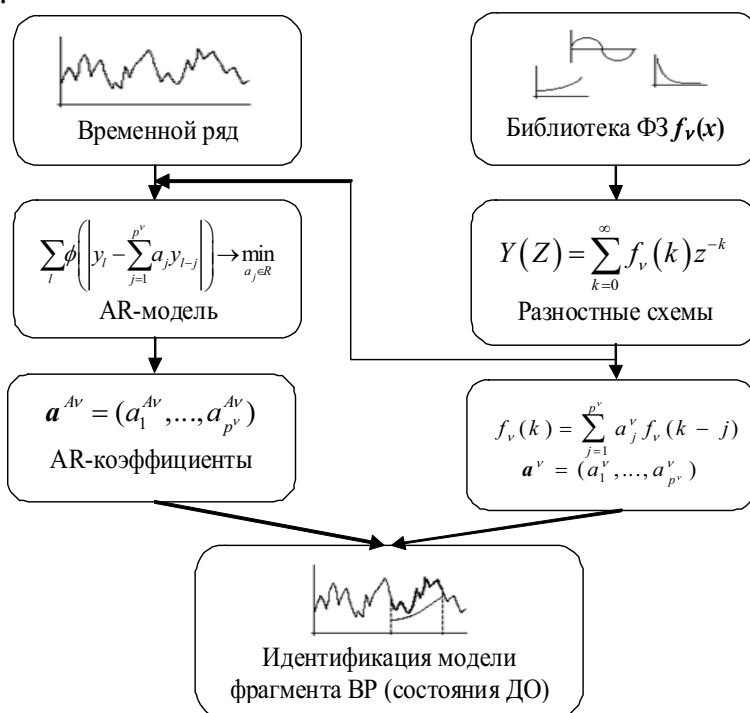


Рис. 2.5. Концептуальная схема идентификации моделей состояний ДО на основе совместного применения разностных схем и AR-моделей функциональных зависимостей (ФЗ) из заданного класса, ОДЗ – область допустимых значений

В подходе, являющимся основой алгоритма 2.5, устанавливается соответствие между классами функциональных зависимостей, которым однозначно отвечают коэффициенты разностных схем (рис. 2.5), образующих центры областей допустимых значений (ОДЗ), и коэффициентами авторегрессионной зависимости (AR-модели), построенной по g -эталонам временного ряда.

Такой подход развивался в работах Семёнычева В.К. [147] для моделирования и прогнозирования экономических рядов динамики сезонных и циклических компонент, а также в работах А.Н. Тырсина. Ниже излагается модификация этого подхода применением не обычного равномерного усреднения данных регрессии при группировании значений, а неравномерного группирования посредством g -эталонов, и его применение для решения задачи распознавания состояний ДО.

Далее метод моделирования состояний ДО на основе обобщенных эталонов и параметризации модели авторегрессии и соответствующий алгоритм будем называть методом (алгоритмом) ARADS, аббревиатура которого образована от слов AutoRegression, Adaptive algorithm, Difference Scheme, поскольку соответствующий метод основан на совмещении авторегрессионных моделей, разностных схем и распознавания образов и является в настоящем изложении адаптивной моделью по наблюдениям.

Приведем теоретическое обоснование излагаемого ниже алгоритма 2.5 реконструкции временного ряда на основе метода ARADS.

Постановка задачи. Пусть L - линейное нормированное пространство всевозможных числовых последовательностей. Пусть имеется n значений наблюдаемого ВР $y=(y_1, \dots, y_n)$, представляющих собой аддитивную смесь некоторой детерминированной составляющей x_k и белого шума ξ_k с нулевым математическим ожиданием и дисперсией σ_ξ^2 :

$$y_k = x_k + \xi_k, k \geq 0, \quad (2.1)$$

предполагая, что $x_k = f_k = f(k\Delta)$, $\Delta > 0$, где $f(t)$ - некоторая функция, описывающая поведение неслучайной составляющей x_k .

Поставим задачу структурной идентификации неслучайной составляющей x_k на основе заданного множества $f_\nu, \nu = \overline{1, n_f}$ возможных функциональных зависимостей, принадлежащих классу функций, Z -преобразование которых есть дробно-рациональная функция.

Решение задачи. Для решения задачи будем использовать подход, связанный с параметризацией моделей авторегрессии, идейно восходящий к работе [25] (и далее развиваемый в работах [147, 157]), где было высказано предположение о существовании связи между порядком, оценками коэффициентов авторегрессии и типами, параметрами функциональных зависимостей моделируемого ряда динамики. Общая схема варианта решения представлена по шагам.

1. Сопоставляем функциональной зависимости $f_\nu(t), \nu = \overline{1, n_f}$ на основе Z -преобразования разностную схему:

$$x_k^\nu = \sum_{j=1}^{p^\nu} a_j^\nu x_{k-j}^\nu, \quad \forall k \geq p^\nu, \quad x_k^\nu = f_\nu(t_k) = f_\nu(k\Delta), \quad \Delta > 0. \quad (2.2)$$

Сопоставление (2.2) всегда возможно для класса функций $f_\nu(t)$ с дробно-рациональным Z -преобразованием или приводящих к нему определенной заменой переменных $\varphi_\nu(f_\nu(t))$, например, линейризацией (см. ниже табл. 2.4), одновременно решающей проблему минимизации числа коэффициентов AR-моделей разных трендовых моделей (функциональных зависимостей). Во втором случае выбор лучшей модели из заданного множества кривых возможен, но с отдельной задачей параметрической идентификации.

В результате получаем набор векторов размерности $\mathbf{a}^\nu(p^\nu) = (a_1^\nu, \dots, a_{p^\nu}^\nu)$, полагаем $p = \max_\nu p^\nu$, $\mathbf{a}^\nu(p) = (a_1^\nu, \dots, a_{p^\nu}^\nu, a_{p^\nu+1}^\nu, \dots, a_p^\nu)$, $a_j^\nu = 0, \forall j = \overline{p^\nu+1, p}$. Значения $\mathbf{a}^\nu(p)$ являются центрами областей D_ν допустимых значений коэффициентов (ОДЗ) ν -й модели, размер которой зависит от точности выбора модели.

2. Выдвигаем n_f предположений относительно вида функциональной зависимости неслучайной составляющей $x_k = f_k$ в модели (2.1) по наблюдаемому ряду (фрагменту). Для проверки n_f гипотез будем строить n_f AR-моделей. Для этого выбираем число точек m , зависящее от заданной пороговой точности аппроксимации и порядка разностной схемы, полученной в п. 1. Известна рекомендация о том [33], что приемлемые результаты по точности в авторегрессионных моделях дает выборка значений в 3–4 раза большая порядка авторегрессии. Например, для параметризации экспоненциальной модели ($y_k = a_1 y_{k-1}, p=1$) достаточно 3-4-х наблюдений. Пусть $m = \max_\nu n_\nu$, где n_ν - число точек, требуемое для построения ν -й модели, $n_\nu > 3p^\nu$, уточняемое на обучающей выборке.

3. Осуществляем построение n_f AR-моделей по выбранным значениям наблюдений $\{y'_j\}$, $y'_j \in (y_1, \dots, y_n)$ ($y'_j \in (z_1, \dots, z_n), z_k = \varphi_\nu(y'_k)$):

$$(a_1^{A\nu}, \dots, a_{p^\nu}^{A\nu}) = \arg \min_{a_j^\nu \in R} \sum_{l \in \{1, \dots, m\}} \phi \left(\left| y'_l - \sum_{j=1}^{p^\nu} a_j^\nu y'_{l-j} \right| \right), \quad \nu = \overline{1, n_f}, \quad (2.3)$$

$$\text{или } (a_1^{A\nu}, \dots, a_{p^\nu}^{A\nu}) = \arg \min_{a_j^\nu \in R} \sum_{l \in \{1, \dots, m\}} \phi \left(\left| z_l^\nu - \sum_{j=1}^{p^\nu} a_j^\nu z_{l-j}^\nu \right| \right), \quad \nu = \overline{1, n_f}.$$

где ϕ - некоторая монотонно возрастающая на $(0, \infty)$ функция, $\phi(0)=0$, $\phi'(x) < 0$ для любого $x > 0$, $\sup_{x \in R^+} |x^2 \phi'(x)| < \infty$.

4. Определяем тип функциональной зависимости f_{v_0} в смысле близости векторов $\mathbf{a}^{Av}(p)$ и $\mathbf{a}^v(p)$ по заданной норме $v_0 = \arg \min_{v=1, n_f} \|\mathbf{a}^v - \mathbf{a}^{Av}\|$, например:

$$v_0 = \arg \min_{v=1, n_f} \sum_{v=1}^p r(\mathbf{a}^v, \mathbf{a}^{Av}), r(\mathbf{a}^v, \mathbf{a}^{Av}(i)) = \sqrt{\sum_{j=1}^p (a_j^v - a_j^{Av})^2}. \quad (2.4)$$

Подытожим вышесказанное в виде утверждений 1-3, доказательство которых приведено в приложении 2.

Утверждение 2.1. Пусть выполнены условия:

1) элементы последовательности наблюдений $\mathbf{y}=(y_0, y_1, \dots)$ подчиняются закону: $y_k = x_k + \xi_k$, где ξ_k - белый шум с нулевым математическим ожиданием и дисперсией σ_ξ^2 ; x_k - значение некоторой равномерно дискретизованной функции: $x_k = f(k\Delta; \alpha_1, \alpha_2, \dots, \alpha_m)$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)$ - вектор параметров, $\mathbf{x} = \{x_k\}$, $k \geq 0$;

2) функция $x_k = f(k\Delta)$ принадлежит классу функций, Z-преобразование которых представимо в виде дробно-рациональной функции:

$$X(z) = \sum_{j=0}^{\infty} x_j z^{-j} = P(z)/Q(z), R(z) = \sum_{j=0}^{p'} p_j z^{-j}, Q(z) = \sum_{j=0}^{q'} q_j z^{-j}, p' \leq q' < \infty. \quad (2.5)$$

Тогда последовательности \mathbf{x} однозначно соответствует разностная схема

$$x_n = \sum_{j=1}^p a_j x_{n-j}, \quad \forall n \geq p, \text{ а последовательность } \mathbf{y} \text{ представляет собой процесс авторегрессии - скользящего среднего ARMA}(p, p).$$

Выбор «лучшей» модели из любого заданного множества моделей возможен на основании утверждения 2.2.

Утверждение 2.2. Пусть имеем последовательность функций $\{f(k)\}$ ($k=1, 2, \dots$) и заданное произвольное множество $F = \{f_j(k)\}$, $j \in J$ (J - некоторое множество индексов) непрерывных, ограниченных снизу функций одной переменной. Тогда множество F упорядочено по отношению порядка в смысле близости по норме к заданной функции f .

Известно¹⁵, что процедура вычисления коэффициентов авторегрессии на основе МНК и метода наименьших модулей неустойчива к выбросам.

Утверждение 2.3. Оценки $(a_1^{Av}, \dots, a_p^{Av})$ AR-моделей, полученные по формуле (2.3) с монотонно возрастающей на $(0, \infty)$ функцией потерь ϕ , обладающей свойствами: $\phi(0)=0$, $\phi'(x) < 0$ для любого $x > 0$, являются устойчивыми к большим ошибкам.

Пример 2.2. Пусть исследуется зависимость $x = A \exp(\alpha t)$, соответствующая последовательность $x_k = A \exp(\alpha t_k) = A \exp(\alpha \Delta k)$, k - номер отсчета.

¹⁵ Болдин М.Б., Симонова Г.И., Тюрин Ю.Н. Знаковый статистический анализ линейных моделей. - М.: ФИЗМАТЛИТ, 1997. - 288 с.

Применяем к последовательности отсчетов $x_k = A \exp(\alpha \Delta k)$ Z -преобразование, получим в области изображений:

$$X(Z) = \sum_{k=0}^{\infty} A \exp(\alpha \Delta k) z^{-k} = A \sum_{k=0}^{\infty} \exp(k(\alpha \Delta - \ln z)) = A(1 - \exp(\alpha \Delta - \ln z))^{-1} = A(1 - \exp(\alpha \Delta) z^{-1})^{-1},$$

$$X(Z)(1 - \exp(\alpha \Delta) z^{-1}) = A. \quad (2.6)$$

Для последнего уравнения выполняем обратное Z -преобразование, в области оригиналов получим разностную схему: $x_k = \exp(\alpha \Delta) x_{k-1} + A \delta(k)$, где $\delta(k)$ – символ Кронекера, откуда следует $x_k = \exp(\alpha \Delta) x_{k-1}$, $k \geq 2$, т.к. $\delta(k) = 0$, $k \geq 2$. Осуществив замену переменной $z_k = \ln x_k$, получим уравнение $z_k = 2z_{k-1} - z_{k-2}$ с центром ОДЗ в 2-мерном пространстве $(a_1, a_2) = (2, -1)$.

Аналогично, для зависимости $y = A + Bt + Ct^2$ получим разностную схему в виде: $x_k = 3x_{k-1} - 3x_{k-2} + x_{k-3}$. Используя замену переменных $z_k = x_k - x_{k-1}$, получаем разностное уравнение вида $z_k = 2z_{k-1} - z_{k-2} = a_1 z_{k-1} + a_2 z_{k-2}$, т.е. центр ОДЗ в 2-мерном базисном пространстве $(a_1, a_2) = (2, -1)$. В данном случае разностная схема представляет собой процесс авторегрессии порядка 2 с нулевым шумом (отсутствие ошибок измерения).

Приведем пример в пользу гипотезы о том, что тип функции, моделирующей временной ряд (2.1), определяет порядок и диапазон значений коэффициентов авторегрессии [25, 147].

Пример 2.3. Пусть в (2.1) моделирование тренда x_k производится на основе 2-х моделей функциональных зависимостей:

$$f_k^1 = a_1 \exp(-\alpha_1 t_k) + a_2 \exp(-\alpha_2 t_k), f_k^2 = \exp(-\alpha_3 t_k)(a_3 t_k + a_4) \quad (2.7)$$

с соответствующими авторегрессиями 2-го порядка:

$$y_k^1 = \lambda_1^1 y_{k-1}^1 - \lambda_2^1 y_{k-2}^1 + \xi_k, y_k^2 = \lambda_1^2 y_{k-1}^2 - \lambda_2^2 y_{k-2}^2 + \xi_k,$$

где коэффициенты $\lambda_1^1 = \exp(-\alpha_1 \Delta) + \exp(-\alpha_2 \Delta)$, $\lambda_2^1 = \exp(-(\alpha_1 + \alpha_2) \Delta)$ - для модели f^1 , и $\lambda_1^2 = 2 \exp(-\alpha_3 \Delta)$, $\lambda_2^2 = \exp(-2\alpha_3 \Delta) / 2$ - для модели f^2 .

Анализ коэффициентов авторегрессии показывает, что на данном этапе идентификации при $\alpha_1 > 0$, $\alpha_2 > 0$ к условиям отнесения временного ряда (фрагмента) к первой модели будет система неравенств: $0 < \lambda_1 < 2$, $0 < \lambda_2 < 0.25 \lambda_1^2$, ко второй модели: $0 < \lambda_1 < 2$, $\lambda_2 = 0.25 \lambda_1^2$.

Реализация МНК приводит к уравнению для оценок $\hat{\lambda}_1, \hat{\lambda}_2$ параметров λ_1, λ_2 :

$$(\hat{\lambda}_1, \hat{\lambda}_2) = \arg \min_{\lambda_1, \lambda_2} \sum_k (y_k - \lambda_1 y_{k-1} - \lambda_2 y_{k-2})^2, \quad (2.8)$$

далее к «нормальной» системе алгебраических линейных уравнений 2-го порядка, решения которой определяют параметры моделей (2.7):

$$\hat{\alpha}_{1,2} = -\frac{1}{\Delta} \ln \left(\frac{\hat{\lambda}_1}{2} + \left(\frac{\hat{\lambda}_1^2}{4} \pm \hat{\lambda}_2 \right)^{1/2} \right); \hat{\alpha}_3 = -\frac{1}{\Delta} \ln \frac{\hat{\lambda}_1}{2}.$$

Далее найденные оценки $\hat{\alpha}_{1,2}$, $\hat{\alpha}_3$ используют для нахождения оценок параметров a_1 , a_2 (или a_3 , a_4), входящих в обе модели (подставляя их в соответствующую «нормальную» систему).

Пример 2.4. Приведем пример соответствий [157] некоторых функциональных зависимостей и моделей разностных схем (табл. 2.4).

Таблица 2.4. Примеры соответствий коэффициентов разностных схем кривым

№	Модель тренда	Замена переменной $z_k = \varphi_\nu(y_k)$	Разностная схема	Центр ОДЗ в пространстве коэффициентов
1	$y_k = A + Bt_k$	-	$y_k = 2y_{k-1} - y_{k-2}$	$(a_{y1}, a_{y2}) = (2, -1)$
2	$y_k = A + Bt_k + Ct_k^2$, $y_k = 3y_{k-1} - 3y_{k-2} + y_{k-3}$	$z_k = y_k - y_{k-1}$	$z_k = 2z_{k-1} - z_{k-2}$	$(a_{z1}, a_{z2}) = (2, -1)$ $(a_{y1}, a_{y2}, a_{y3}) = (3, -3, 1)$
3	$y_k = A + Bt_k^{-1}$	$z_k = y_k t_k$	$z_k = 2z_{k-1} - z_{k-2}$	$(a_{y1}, a_{y2}) = (2, -1)$
4	$y_k = A \exp(Bt_k)$	$z_k = \ln y_k$	$z_k = 2z_{k-1} - z_{k-2}$	$(a_{y1}, a_{y2}) = (2, -1)$

Комментарий к таблице 2.4. Переход к новым обозначениям (замена переменных) связан с минимизацией числа AR-коэффициентов, определяющих размер пространства, в котором находится область допустимых значений, идентифицирующая модель тренда. В случае осуществления замены переменных $z_k = \varphi_\nu(y_k)$ при получении разностной схемы построение авторегрессионной модели должно производиться не по наблюдениям y_k , а по новому временному ряду $z_k = \varphi_\nu(y_k)$.

Алгоритм 2.5. Реконструкция состояний ДО на основе g-эталонов и метода ARADS

Вход: множество g-эталонов $U_G = \{G_1, G_2, \dots, G_m\}$, с координатами j-го g-эталона: (y_j, t_j') , $y_j = \overline{Y}(t_j')$, $j = \overline{1, m}$, и ВК w_j ; таблица соответствий видов функциональных зависимостей и ОДЗ коэффициентов a_1^ν, \dots, a_p^ν , $\nu = \overline{1, n_f}$ разностных схем (табл. 2.4); ε - точность подгонки параметров функциональной модели.

Выход: признаки состояний ДО: ОДЗ коэффициентов разностной схемы $D_i = (a_1^{v_i}, \dots, a_p^{v_i})$, соотнесенной с i-м состоянием; признак формы ДО $f_{v_i}, i = \overline{1, I}$.

1: Выделяем на интервале длительностью T_i , соотнесенном состоянию Ω_i , κ_i равномерно расположенных точек ($\kappa_i > 2$): $t_{i0}'' = t_i, \dots, t_{il}'' = t_i + lh_i, \dots, t_{i\kappa_i}'' = t_{i+1}$, где $T_i = t_{i+1} - t_i$, $h_i = T_i / \kappa_i$.

2: Сопоставляем каждой l -й точке t_{il}'' ($l \in \{1, \dots, \kappa_i\}$) усредненный g_a -эталон, значение ординаты которого равно: $y_l' = \nu_l^{-1} \sum_{e \in V} y(t_e'')$, где V_l – множество g -эталонов, уложившихся в интервал $(t_{il}'', t_{i,l+1}'')$, $\nu_l = |V_l|$ – мощность множества V_l (число g -эталонов в l -м интервале); если $\forall j = \overline{1, m} \quad t_j' \notin (t_{il}'', t_{i,l+1}'')$, (в интервал (t_l', t_{l+1}') не попало ни одного g -эталона), то в качестве ординаты и абсциссы l -го g_a -эталона полагаем значения ординаты и абсциссы предыдущего g -эталона.

3: Полагаем значения g_a -эталонов $\{y_1', \dots, y_{\kappa_i}'\}$ (или $z_1', \dots, z_{\kappa_i}'$, $z_k' = \varphi_\nu(y_k)$) в качестве исходных данных для построения модели $AR(p^\nu)$ для заданных классов моделей функциональных зависимостей $f_\nu, \nu = \overline{1, n_f}$ (n_f – мощность класса) по значениям временного ряда, фрагмент которого соотнесен с Ω_i :

$$\left(a_1^{A\nu}(i), \dots, a_{p^\nu}^{A\nu}(i) \right) = \arg \min_{a_j^i \in R} \sum_{l \in \{1, \dots, \kappa_i\}} \phi \left(\left| y_l' - \sum_{j=1}^{p^\nu} a_j^i y_{l-j}' \right| \right), \quad i = \overline{1, I}, \quad (2.9)$$

где ϕ – некоторая монотонно возрастающая на $(0, \infty)$ функция, $\phi(0) = 0$, $\phi'(x) < 0$ для любого $x > 0$, $\sup_{x \in R^+} |x^2 \phi'(x)| < \infty$.

4: Определяем признак i -го состояния как ОДЗ $D_i = (a_1^{v_i}, \dots, a_{p_i}^{v_i})$, $i = \overline{1, I}$ коэффициентов разностной схемы, отвечающей наиболее близкой модели f_{v_i} :

$$v_i = \arg \min_{\nu=1, n_f} \sum_{j=1}^p r(a^\nu, a^{A\nu}(i)), \quad r(a^\nu, a^{A\nu}(i)) = \sqrt{\sum_{j=1}^p (a_j^\nu - a_j^{A\nu})^2}, \quad p = \max_\nu p^\nu, \quad (2.10)$$

и уравнение кривой состояния ДО (признак формы) $f_{v_i}, i = \overline{1, I}$.

5: Осуществляем подгонку параметров выбранной модели, решая условную систему алгебраических неравенств: $|f_{v_i}(t, \mathbf{b}) - y(t)| < \varepsilon$, $t \in \{1, \dots, \kappa_i\}$, где \mathbf{b} – вектор параметров модели.

Вернуть $D_i = (a_1^{v_i}, \dots, a_{p_i}^{v_i})$, $f_{v_i}, i = \overline{1, I}$.

Замечания.

1. Изложенный эвристический алгоритм 2.5 формирует признаки формы состояний ДО, не зависящие от значений параметров кривой, то есть, по сути, реализует непараметрический подход. Однако, для полноты обсуждаемой модели и уверенности в применении к конкретному временному ряду (выборке), необходимо проведение дополнительных исследований, связанных с проверкой не

только значимости получаемых коэффициентов разностных схем как признаков состояний ДО, но и взаимной некоррелированности случайных остатков $y(t) - f_{v_i}(t, \mathbf{b}), y(t) \in \Omega_i$

2. Как известно [25, 28], в случае статистической значимости нескольких регрессионных моделей ввиду разной формы их функциональной зависимости выбор наилучшей оказывается весьма затруднительным и, как правило, разрешается только экспертно. Разностная схема вышеприведенного алгоритма 2.5 позволяет это осуществить в силу статистической эквивалентности разностной схемы и процессов авторегрессии – скользящего среднего порядка ARMA(p, p) (утверждения 2.1, 2.2).

3. На основе ARADS возможна идентификация связи между двумя показателями. Напомним, что к существенной особенности постановок задач (1.1) и (1.2) относится принципиальная неизмеряемость части координат. Для идентификации сложного динамического объекта необходимо по наблюдаемым данным восстанавливать ее входные неизмеряемые координаты (в частности, строить наблюдатель в сложных технических системах для «скрытых» переменных).

Пусть получен некоторый объем обучающей выборки с известными сочетаниями показаний наблюдаемой Y и ненаблюдаемой X переменных (x_j, y_j) , $j = \overline{1, n}$, характеризующих анализируемые параметры X и Y некоторого объекта в моменты времени $t=j\Delta$. Задача будет заключаться в нахождении зависимости $X=\varphi(Y)$.

Способ идентификации связи между двумя показателями на основе ARADS (алгоритмы 2.5, 2.6) заключается в следующем.

Для каждого из ВР X и Y строятся (идентифицируются) динамические модели по алгоритму 2.5: $x=g(t)$, $y=f(t)$, в результате получаем регрессионную модель между наблюдаемой (независимой) Y и ненаблюдаемой (объясняющей) X переменными: $x = g(f^{-1}(t))$, где $f^{-1}(\cdot)$ - функция, обратная к $f(\cdot)$. На контрольной выборке проверяется значимость полученной на обучающей выборке зависимости.

Пример 2.5. Приведем примеры функций потерь с указанными свойствами и устойчивыми к большим выбросам:

$$\phi(x) = |x|/(|x| + 1), \phi(x) = \arctg(|x|), \phi(x) = 1 - \exp(-|x|). \quad (2.11)$$

Примерами функций потерь с неограниченной чувствительностью к большим выбросам являются следующие:

$$\phi(x) = |x|^\lambda, \lambda > 0, \phi(x) = \ln(|x| + 1).$$

На основе метода моделирования состояний с применением вышеизложенного способа параметризации модели авторегрессии возможно решение

важных прикладных задач, связанных с повышением надежности распознавания состояний ДО и качества цифрового дифференцирования, примененного для построения наблюдателя при управлении асинхронным двигателем (АД), гл. 7, 9.

Идентификация пространственной зависимости (формы состояний) осуществляется по алгоритму 2.6.

Алгоритм 2.6. Идентификация состояний ДО по признаку формы

Вход: множество наблюдаемых значений (k -окно): (y_j, t_j) , $y_j = Y(t_j)$, $j = \overline{1, k}$,

где \bar{y}_j – усредненное значение глубины k : $\bar{y}_j = \frac{1}{k} \sum_{e=t_j-k+1}^{t_j} y_e$; множество призна-

ков состояний ДО: ОДЗ в виде коэффициентов разностной схемы $D_i = (a_1^{v_i}, \dots, a_p^{v_i})$, соотнесенной с v_i -й функциональной зависимостью, сопоставленной состоянию Ω_i ; уравнения кривых состояний ДО (признаков формы) f_{v_i} , $i = \overline{1, I}$.

Выход: метка (номер) i_0 состояния ДО, которому принадлежит наблюдаемое k -окно.

1: Строим n_f AR-моделей (для каждой кривой f_v , $v = \overline{1, n_f}$) по выборке (t_j, \bar{y}_j) , определяем коэффициенты AR-моделей $(a_1^{Av}, \dots, a_{p^v}^{Av})$, $v = \overline{1, n_f}$:

$$(a_1^{Av}, \dots, a_{p^v}^{Av}) = \arg \min_{a_j^{Av} \in R} \sum_{l \in \{1, \dots, k\}} \phi \left(y_l - \sum_{j=1}^{p^v} a_j^{Av} y_{l-j} \right),$$

где ϕ вышеопределенная функция (утверждение 2.3)

2: Находим расстояние $r(a^{v_i}, a^{Av})$ между векторами $(a_1^{v_i}, \dots, a_p^{v_i})$, $(a_1^{Av}, \dots, a_{p^v}^{Av})$: $r(a^{v_i}, a^{Av}) = \sqrt{\sum_{j=1}^{\max(p^v, p)} (a_j^{v_i} - a_j^{Av})^2}$ в p -мерном базисном пространстве,

координатами которого служат коэффициенты разностной схемы.

3: Принимаем решение в пользу i_0 -го состояния: $i_0 = \arg \min_i r(a^{v_i}, a^{Av})$.

Вернуть значение i_0 .

2.4. Применение алгебраического подхода к анализу зашумленных временных рядов

В работе [139] изложен алгебраический подход к задаче построения обучаемых алгоритмов выделения трендов. В его основе лежит идея разметки точек временного ряда аксиомами (условиями).

Аксиома – это некоторое правило, которому должна удовлетворять одна или более точек ряда. Если набор точек $\{y_i\}$ удовлетворяет аксиоме A , то будем говорить, что аксиома A выполняется для этих точек. Пример аксиомы [44]: «аксиома задается на $(k-j)$ точках; точка x_i является точкой минимума для интервала $[x_j, x_k]$, если $x_i < x_l, \forall l \in [j, k]$ ».

Анализ немногих работ в этом направлении [44, 60, 139] и исследование соответствующих алгоритмов показали, что на данный момент существуют две нерешенные проблемы, ограничивающие практическое применение алгебраического подхода: во-первых, системы аксиом создаются либо вручную на основе априорных данных (что для длинных ВР неприемлемо), либо на основе генетического алгоритма с открытыми вопросами о стартовых условиях и сходимости; во-вторых, отсутствие общих рекомендаций для разметки стохастических ВР. На ряде модельных примеров¹⁶ показано, что доля верно распознанных образов с ростом показателя γ , равному отношению максимальной амплитуды шума к максимальной амплитуде сигнала, постепенно снижается (с 96.1% при $\gamma=0.025$ до 66.8% и ниже при $\gamma=0.25$).

В данном разделе предлагается метод к применению алгебраического подхода на случай зашумленных ВР с автоматической разметкой ВР и принятием решения о метке (номере) состояния ДО по окну наблюдения. Уточним постановку задачи для изложения метода.

Постановка задачи распознавания образов на временных рядах. Рассмотрим временной ряд $\{y_i\}$, сопровождающий функционирование некоторого ДО и содержащий характерные фрагменты (области) нескольких видов (здесь и далее, состояний ДО, или мета-образов). Будем предполагать, что по совокупности признаковов описания состояния ДО (мета-образы) представляют собой последовательности одного и более элементарных образов (из конечного домена U_{is}), которые отличаются между собой хотя бы по одному элементарному образу.

Элементарный образ определим как некоторое множество «близких» траекторий в фазовом пространстве ДО, устойчивых образцов поведения (с точки зрения наблюдателя), демонстрируемых системой в сходных условиях.

Домен U_{is} составлен из эталонных или реально наблюдавшихся в прошлом рядов, содержащих образы, с указанием типов образов и их границ.

Задача распознавания образов заключается, во-первых, в выделении фрагментов, содержащих (искаженные шумом и нелинейностью ДО) образы; во-вторых, в определении типов возникающих образов; в-третьих, в идентификации подпоследовательностей элементарных образов, близких к эталонным, определяющих метку состояния ДО (мета-образа).

¹⁶ Коваленко Д.С., Костенко В.А., Васин Е.А. Исследование применимости алгебраического подхода к анализу временных рядов // Тезисы конференции МСО-2005.

Метод решения задачи. Метод (следуя [44, 60, 139]) включает в себя шаги: 1) построение системы аксиом; 2) установление соответствия между каждым выделенным (описанием характеристик, или признаков) состоянием ДО (мета-образом) и «эталонной» последовательностью аксиом; 3) автоматическая разметка наблюдаемого ВР выбранной системой аксиом; 4) распознавание, или поиск подпоследовательностей, близких к эталонным. Первые два шага составляют этап обучения, следующие два - этап распознавания.

Целью этапа обучения является построение системы аксиом, «хорошо» описывающих представленные образы, а именно, система аксиом должна обладать, с одной стороны, свойством обобщения для отыскания зашумленных и искаженных образов во временной последовательности; с другой стороны, дискриминирующим свойством для различения образов между собой.

Целью этапа распознавания является построение последовательности аксиом, которым удовлетворяют точки фрагментов анализируемого ряда (состояний ДО).

Пример 2.6. Приведем пример системы аксиом, удовлетворяющей свойствам полноты и однозначности [44].

Обозначим: $x_1 = x_{i-k}$, $x_2 = x_{i+j}$, $t_1 = i-k$, $t_2 = i+j$, $t = l$. Пусть функция $f(t)$ задает прямую, соединяющую точки $x_1 = x_{i-k}$, $x_2 = x_{i+j}$:

$$f(t) = \frac{x_2 - x_1}{t_2 - t_1} t + \frac{t_2 x_1 - t_1 x_2}{t_2 - t_1} = \frac{x_{i+j} - x_{i-k}}{k+j} l + \frac{k-j}{k+j} x_{i+j} + \frac{i+j}{k+j} x_{i-k} = f(l).$$

Параметры m и $M > m$, $m \leq k+j \leq M$ определяют минимальный и максимальный размер участка ряда, который может быть размечен одной аксиомой. Точка ряда x_i размечается аксиомой A , если в этой точке выполняется условие аксиомы. Система аксиом, обладающая указанными качествами, имеет следующее описание:

$A_1 : \forall l \in (1, k+j), l \in N : x_{l-1} < x_l < x_{l+1}, f(l) - x_l > \delta$ (возрастание функции, возрастание первой производной);

$A_2 : \forall l \in (1, k+j), l \in N : x_{l-1} < x_l < x_{l+1}, f(l) - x_l < -\delta$ (возрастание функции, убывание первой производной);

$A_3 : \forall l \in (1, k+j), l \in N : x_{l-1} > x_l > x_{l+1}, f(l) - x_l < -\delta$ (убывание функции, убывание первой производной);

$A_4 : \forall l \in (1, k+j), l \in N : x_{l-1} > x_l > x_{l+1}, f(l) - x_l > \delta$ (убывание функции, возрастание первой производной);

$A_5 : \forall l \in (1, k+j), l \in N : x_{l-1} < x_l < x_{l+1}, \exists l \in (1, k+j)$, для которых имеет место невыполнение аксиом A_1 и A_2 (возрастание функции, вторая производная меняет знак);

$A_6 : \forall l \in (1, k + j), l \in N : x_{l-1} > x_l > x_{l+1}, \exists l \in (1, k + j)$, для которых имеет место невыполнение аксиом A_3 и A_4 (убывание функции, вторая производная меняет знак);

A_0 : не выполнение аксиом A_1, \dots, A_6 . Значение параметра δ определяет допустимую погрешность невыполнения аксиом.

Зададим следующее правило разметки ВР в постановке (2.1), где неслучайная составляющая описывает скачкообразное поведение моделируемого динамического процесса, сопоставленное смене состояний ДО.

Аксиома A_{v_0} задается на $(k''-k')$ точках; точка $x_l \forall l \in [k', k'']$ является точкой с меткой f_{v_0} для интервала $[x_{k'}, x_{k''}]$, если имеет место (2.4), где $\mathbf{a}^{Av} = \mathbf{a}^{Av}[k', k'']$ - вектор AR-коэффициентов, восстановленный по интервалу $[k', k'']$, \mathbf{a}^v - коэффициенты разностной схемы, соответствующей функции f_v .

Алгоритм 2.7 разметки ВР на основе данной системы аксиом и решения на его базе задачи распознавания состояний ДО приведен описательно как решение отдельных подзадач.

Алгоритм 2.7. Разметка временных рядов и распознавание состояний

Вход: Ряд: y_1, \dots, y_N (набор рядов $\mathbf{y}_j = \{y_{j1}, \dots, y_{jN}\}, j = \overline{1, J}$ в задаче сравнения и классификации); геометрические границы образов состояний ДО, параметры алгоритма: величины скользящего окна ($k \geq k_0$) и сдвига k -окна (здесь для простоты изложения, равному k); эталонные размеченные подпоследовательности, сопоставленные разным состояниям ДО.

Выход: размеченный ВР: $l_f = (f_{j_1}, \dots, f_{j_N}), f_j \in \{1, \dots, n_f\}$.

1: Разметка ряда по значениям. Сопоставим последнему значению скользящего k -окна метки f_j функции из заданного алфавита F по алгоритму 2.5.

Вернуть: $l_f = (f_{j_1}, \dots, f_{j_{N-k}}), f_j \in \{1, \dots, n_f\}$.

2: Вычисление меры сходства между ВР $\mathbf{y}_j, \mathbf{y}_k$. Поставим ВР \mathbf{y}_j в соответствие: а) ряд $\mathbf{d}_j^v = \{d_{jt}^v\}, d_{jt}^v = 1$, если значению y_{jt} присвоена метка f_v и $d_{jt}^v = 0$ иначе ($t = \overline{1, N}$); б) набор интервалов $\Lambda_{j\xi}^v$, состоящих из последовательно идущих меток f_v , где ξ - номера вхождений подпоследовательностей из меток f_v . Тогда меру сходства r_{jk} между ВР $\mathbf{y}_j, \mathbf{y}_k$ можно определить как число «похожих» интервалов с параметром точности δ [60]: $r_{jk} = \sum_{v=1}^{n_f} r_{jk}^v$, где

$$r_{jk}^v = \sum_{(\xi, \zeta)} \chi_{jk}^v(M_{\xi\zeta}), M_{\xi\zeta} = \{|\Lambda_{j\xi}^v \cap \Lambda_{k\zeta}^v| + \delta \geq |\Lambda_{j\xi}^v \cup \Lambda_{k\zeta}^v|\}, \chi_{jk}^v(M_{\xi\zeta}) -$$

индикатор события $M_{\xi\xi}, |X|$ - мощность множества X .

3: Классификация рядов. Поставим матрице смежности $S = \|s_{jk}\|, s_{jk} = \begin{cases} 1, r_{jk} \geq r_0 \\ 0, r_{jk} < r_0 \end{cases}$ в соответствие полносвязный граф, анализ которого

позволит выделить «похожие» объекты [151], описываемые временными рядами.

4: Разметка ряда по состояниям ДО производится на основе поиска моментов вхождения наиболее «близких» заданных эталонных отрезков с использованием введенной выше меры сходства.

Вернуть: размеченный ВР (ряды), классы ВР, метка состояния ДО при наблюдении фрагмента ВР.

В табл. 2.5 приведены результаты работы 2-х приведенных алгоритмов при различных уровнях шума. В последнем столбце указаны через «/» данные для детерминированной разметки по алгоритму из [44] и на основе ARADS, соответственно.

Таблица 2.5. Характеристики 2-х алгоритмов при разных уровнях шум/сигнал

γ (шум/сигнал)	Доля верных распознаваний при детерминированной разметке, %	Доля верных распознаваний на основе ARADS, %	Максимальное значение доли ложных распознаваний, %
0.0	99.9	100	6.1/0
0.025	96.1	99.1	7.0/0.09
0.05	87.8	98.2	7.3/1.1
0.1	79.1	97.9	7.2/2.1
0.15	75.5	97.4	6.4/2.9
0.2	71.2	96.3	7.1/3.9
0.25	55.5	96.1	7.4/4.6
0.3	45.5	94.9	8.8/5.4
0.35	41.2	94.2	9.1/5.5

Отличительной особенностью метода, реализованного алгоритмом 2.7, является подход к формированию аксиом, который, во-первых, можно автоматизировать, а во-вторых, применять и для зашумленных ВР.

Приведенные данные свидетельствуют о некотором преимуществе данной методики перед существующей [60] и опирающейся на меру сходства между временными рядами, выраженную числом размеченных интервалов совместного роста или спада значений показателей.

2.4.1. Алгоритм реконструкции нестационарного временного ряда на основе обобщенных эталонов и модификации метода Бока

Постановка задачи. Пусть задан временной ряд $y_n, n \geq 1$:

$$y_n = x_n + \zeta_n, x_{n+1} = f(x_n, \theta) + \xi_n, \quad (2.12)$$

где ξ_n - интерпретируется как динамический шум (независимые одинаково распределенные случайные величины), ζ_n - измерительный шум. Ставится задача получения оценки вектора параметров θ для модели (2.12). Рассмотрим модель (2.12) для 3-х случаев.

Случай 1: Пусть $\zeta_n = 0, x_{n+1} = f(x_n, \theta) + \xi_n, y_n = x_n$.

Если распределение $p(\xi)$ - известно, то при достаточно общих условиях наиболее эффективен метод максимального правдоподобия:

$$\ln L(\theta) = \ln p(y_1, y_2, \dots, y_N | \theta) = \sum_{n=1}^{N-1} \ln p_{\xi}(y_{n+1} - f(y_n, \theta)) \quad (2.13)$$

Если ξ_n - нормально распределенный шум, то максимум $\ln L(\theta)$ эквивалентен обычному методу наименьших квадратов:

$$S(\theta) = \sum_{n=1}^{N-1} (y_{n+1} - f(y_n, \theta))^2 \rightarrow \min \quad (2.14)$$

Задачи максимизации максимального правдоподобия (2.13) и минимизации наименьших квадратов (2.14) дают асимптотически несмещенные и состоятельные оценки для стационарного ВР.

Случай 2: Пусть $\xi_n = 0, x_{n+1} = f(x_n, \theta), y_n = x_n + \zeta_n$.

Если ζ_n - нормально распределенный шум, то задача ММП сводится к минимизации суммы квадратов отклонений реализации модели от наблюдаемого ряда:

$$S(\theta, x_1) = \sum_{n=0}^{N-1} (y_{n+1} - f^{(n)}(x_1, \theta))^2 \rightarrow \min \quad (2.15)$$

где $f^{(n)}$ - n -я итерация отображения $x_{n+1} = f^{(n)}(x, \theta), f^{(0)}(x, \theta) = x$, и в число оцениваемых величин включено начальное состояние модели x_1 .

Если ζ_n - шум с неизвестным распределением и большой дисперсией (переменные x_n - скрытые), то задача в постановке (2.12) является открытой. Отдельные способы решения дают смещенные оценки параметров.

Метод Бока [3] представляет собой задачу условной минимизации.

Основной принцип метода заключается в том, что начальные условия $y_1^0(t_1), \dots, y_D^0(t_1)$ (D - размерность ВР) для скрытых (неизмеряемых) переменных включаются в число неизвестных параметров. Для начальных условий задаются стартовые догадки при оценивании параметров $\{\theta_s^{(1)}, \theta_s^{(2)}, \dots, \theta_s^{(p)}\}$ (нижний индекс-

номер стартовой догадки). Для каждой переменной задаётся несколько (L) – начальных условий $y_k^0(t_1), y_k^0(t_{n+1}), \dots, y_k^0(t_{(L-1)n+1})$ ($1 \leq k \leq D$) на разных участках временного ряда. Затем формулируется критерий, на основе которого стремятся достигнуть максимально возможной близости между траекториями наблюдаемых и соответствующих им модельных переменных, для чего используется один из итерационных методов глобальной оптимизации:

$$J(\theta, y_1^0, \dots, y_{(L-1)n+1}^0) = \sum_{i=1}^L \sum_{j=1}^n \left(y_{(i-1)n+j} - g^{(i)}(t^{(i-1)} + j, y_{(i-1)n+j}^0, \theta^{(i)}) \right)^2 \rightarrow \min, \quad (2.16)$$

где $\theta = \{\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(p)}\}$, $g^{(i)}(t, y, \theta^{(i)})$ – выражение, задающее i -ю модель траектории ВР, т.е. состояние модели y в момент времени $t = t^{(i-1)} + j, j = \overline{1, n}$ при начальном условии $y_{(i-1)n}^0$ и параметрах $\theta^{(i)}$.

Дополнительно на траекторию накладывается условие непрерывности траектории (сшивания фрагментов моделей):

$$y^0(t_{in+1}) = g^{(i)}(t_{in+1}, y^0(t_{i(n-1)+1})), \quad i = 1, \dots, (L-1), \quad (2.17)$$

где n – число точек в одном фрагменте ($N = Ln$), $g^{(i)}(t_{in+1}, y^0(t_{i(n-1)+1}))$ – значение вектора состояния модели с начальными условиями $y^0(t_{i(n-1)+1})$.

Перечислим особенности традиционного применения метода Бока:

1) отсутствие правила выбора; числа сегментов равной длительности; стартовых догадок для начальных условий по каждой переменной; отсутствие критериев, по которым можно оценить допуски на ошибки в задании стартовых догадок для искомым параметров по сравнению с их «истинными» значениями, чтобы остаться в области притяжения глобального минимума (2.16);

2) существенное ограничение на длину используемого временного ряда;

3) модели аппроксимирующих функций принадлежат одному классу;

4) аппроксимация производится по всему временному ряду.

Скольльзящая модификация метода Бока. Основные положения:

1) используется идеология скользящего окна, размер которого и величина сдвига окна зависит от положения на временной оси интервала рассмотрения;

2) дается правило выбора вида функциональной зависимости для каждого окна (алгоритм 2.5), параметры которой оцениваются по текущему окну;

3) модели аппроксимирующих функций могут быть неравными на разных сегментах, а длина временного ряда не ограничена;

4) траектории сшиваются в скользящем режиме: в j -м окне строится новая траектория и подгоняется к построенной в $(j-1)$ -м окне, которая не изменяется (полагается окончательной) до момента начала j -го окна. Таким образом, «начальное условие» каждой следующей модели задано.

Последовательность длин окон и величин сдвига задается по обучающей выборке (или, в случае отсутствия такой информации, считаются постоянными как параметры алгоритма). Ниже для простоты изложен алгоритм с величиной сдвига, равной длине предыдущего окна.

Алгоритм 2.8. Модификация метода Бока (скользящая аппроксимация)

Вход: Временной ряд $\{y_1, y_2, \dots, y_N\}$, векторный параметр $\mathbf{n} = (n_1, \dots, n_I)$, где n_i – длина i -го окна.

Выход: аппроксимирующая функция $g(t), t \in [t_0, t_N), t_N = N\Delta$:

$$g(t) = \sum_{i=1}^I \chi_i(t) g^{(i)}(t, y^{(i)}, \theta^{(i)}), \quad (2.18)$$

где $\chi_i(t)=1$, если $t \in [t^{(i-1)}, t^{(i)})$, и $\chi_i(t)=0$, иначе; за $g^{(i)}(t, y^{(i)}, \theta^{(i)})$ обозначено выражение, задающее i -ю модель траектории ВР, при начальном условии $y^{(i)}$ в момент времени t и параметрах $\theta^{(i)}$; множество $t^{(i)} = t^{(i-1)} + n_i\Delta$, $i = \overline{1, I-1}$

1: Осуществляется построение траектории $g^{(i)}(t, y^{(i)}, \theta^{(i)})$ в i -м текущем окне по алгоритму 2.5. Пробными шагами ищутся величины сдвига и длины $(i+1)$ -го окна из минимума (2.19).

2: Осуществляется построение траектории $g^{(i+1)}(t, y^{(i+1)}, \theta^{(i+1)})$ в $(i+1)$ -м текущем окне. Получаем вариант модели с разрывной траекторией на интервале $[t^{(i-1)}, t^{(i+1)})$, состоящей из 2-х не стыкующихся друг с другом фрагментов.

3: Задается условие «сшивания» сегментов $g^{(i)}(t, y^{(i)}, \theta^{(i)}) = y^{(i+1)}$, $t = t^{(i)}$ для получения непрерывной траектории модели на интервале $[t^{(i-1)}, t^{(i+1)})$, и критерий минимизации отклонения по интервалу $[t^{(i)}, t^{(i+1)})$, исключая точки траектории i -го (предыдущего) окна:

$$J_i(y^{(i)}, y^{(i+1)}) = \sum_{j=0}^{n_{i+1}} \left(y(t^{(i)} + j) - g^{(i+1)}(t^{(i)} + j, y^{(i+1)}, \theta^{(i+1)}) \right)^2 \rightarrow \min_{\theta^{(i+1)}}, \quad (2.19)$$

Вернуть: $g(t)$.

Замечание. При наличии обучающей выборки для поиска оптимальных длин окон (и величин сдвигов) задается критерий минимизации отклонения реализации модели ВР от наблюдаемого ВР для всего интервала наблюдения:

$$J(\mathbf{n}, y^{(1)}, y^{(2)}, \dots, y^{(I)}) = \sum_{i=1}^{I-1} J_i(y^{(i)}, y^{(i+1)}) \rightarrow \min_{\mathbf{n}}, \quad \mathbf{n} = (n_1, \dots, n_I). \quad (2.20)$$

$$\text{Вернуть: } g(t) = g^{(i)}(t, y^{(i)}, \theta^{(i)}) = y^{(i+1)}, t \in [t_i, t_{i+1}), i = \overline{1, I-1}$$

На рис. 2.6 представлены результаты работы программы, иллюстрирующие зависимость точности аппроксимации в зависимости от числа выделенных фрагментов (как состояний ДО) ВР – модели ЭМС.

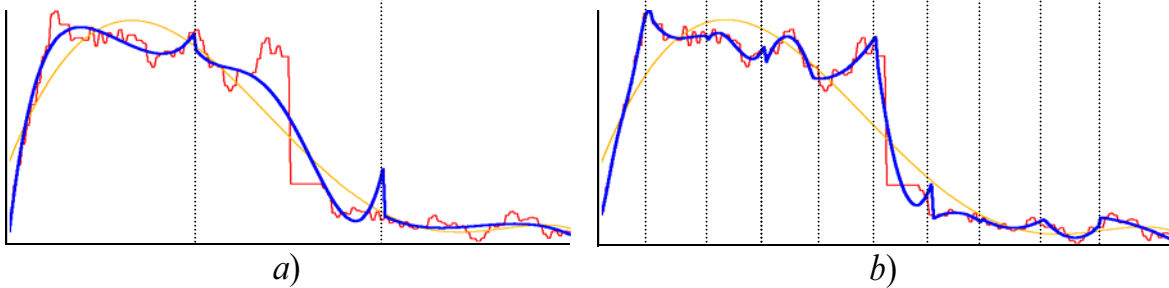


Рис. 2.6. Иллюстрация модификации алгоритма Бока с предварительной аппроксимацией фрагментов ВР (состояний ДО): а) 3 выделенных состояния ДО; б) 10 выделенных состояний ДО

Численные эксперименты на примере нелинейных систем с измерительными шумами показали наибольшую эффективность модификации алгоритма Бока с ослаблением требования непрерывности¹⁷ аппроксимирующей траектории модели на интервале наблюдения. Отметим также, что, несмотря на низкую чувствительность алгоритма к шумам, для коротких сегментов оценка параметров модели может быть смещенной.

2.5. Идентификация состояний ДО на основе непараметрической нелинейной регрессионной модели временного ряда

Систему (1.1) будем рассматривать как стохастический объект, на l входах и скалярном выходе Z которого наблюдаются, соответственно, случайные величины $Y = (Y_1, \dots, Y_l)$ и Z , $Y \in R^l$, $Z \in R^1$, где R^l - l -мерное евклидово пространство.

Положим, что значение выходной переменной Z определяется неизвестным функциональным преобразованием $F(\cdot)$ входных переменных Y , т.е. $Z = F(Y)$. В качестве приближения неизвестной функции $F(\cdot)$ будем использовать функцию регрессии, следуя работе [43]. Функция регрессии или условное математическое ожидание выхода стохастического объекта относительно входов является моделью, которая минимизирует среднеквадратическое отклонение выхода реального объекта и его модели:

¹⁷ V.F. Pisarenko, D. Sornette, Statistical methods of parameter estimation for deterministically chaotic time series. Phys. Rev. E, 2004. V. 69. 036122.

$$r(\mathbf{y}) = \int_{R^1} z f(z | \mathbf{y}) dz = \int_{R^1} z f(\mathbf{y}, z) dz / p(\mathbf{y}) = a(\mathbf{y}) / p(\mathbf{y}), \quad (2.21)$$

где $f(\mathbf{y}, z)$ - неизвестная плотность распределения наблюдаемой $(l+1)$ -мерной случайной величины $(\mathbf{Y}, Z) \in R^{l+1}$ в точке (\mathbf{y}, z) , $f(z | \mathbf{y}) = \frac{f(\mathbf{y}, z)}{p(\mathbf{y})}$ - условная плотность распределения случайной величины $Z \in R^1$ при условии, что $\mathbf{Y} = \mathbf{y}$, $p(\mathbf{y})$ - маргинальная плотность распределения величины \mathbf{Y} , $a(\mathbf{y}) = \int_{R^1} z f(\mathbf{y}, z) dz$.

В качестве непараметрической ядерной оценки условного функционала (2.21) в точке \mathbf{y} в [6] предлагаются полурекуррентные статистики вида:

$$\begin{cases} a_{[n]r}(\mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\prod_{k=1}^l h_{[i]k}} Z_i K\left(\frac{\mathbf{y} - \mathbf{Y}_i}{h_{[i]}}\right) \\ p_{[n]r}(\mathbf{y}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\prod_{k=1}^l h_{[i]k}} K\left(\frac{\mathbf{y} - \mathbf{Y}_i}{h_{[i]}}\right) \end{cases}, \quad r_{[n]}(x) = \frac{a_{[n]r}(\mathbf{y})}{p_{[n]r}(\mathbf{y})}. \quad (2.22)$$

Здесь $a_{[n]r}(\mathbf{y})$, $p_{[n]r}(\mathbf{y})$ - рекуррентные оценки функционала $a(\mathbf{y})$ и плотности $p(\mathbf{y})$, соответственно, (\mathbf{Y}_i, Z_i) , $\mathbf{Y}_i = (Y_{i,1}, \dots, Y_{i,l})$, $i = \overline{1, n}$ - $(l+1)$ -мерная выборка, характеризуемая плотностью $f(x, \mathbf{y})$, $K(\mathbf{u}) = K(u_1, \dots, u_l)$ - l -мерная ядерная функция, $h_{[i]k} > 0$ - последовательность чисел (параметров), сходящаяся к нулю для каждого $k = \overline{1, l}$.

Пример 2.7. Оценка функции авторегрессии. Пусть на выходе объекта наблюдается последовательность одномерных величин $Y_{i+1}, Y_{i+2}, \dots, Y_{i+p+1}$, $i = 1, 2, \dots$, которые образуют процесс авторегрессии порядка p , т.е. связаны соотношением

$$Y_{i+p+1} = \Psi(Y_{i+1}, \dots, Y_{i+p}) + \varepsilon_{i+p+1}, \quad i = 1, 2, \dots, \quad (2.23)$$

где $\{\varepsilon_j\}$ - последовательность независимых и одинаково распределенных случайных величин. Пусть Y_1, \dots, Y_{n+p+1} - выборка объема $n+p+1$, генерируемая процессом (2.23). В [43] показано, что для оценки функции Ψ целесообразна статистика:

$$\Psi_{[n]p}(y) = \sum_{i=1}^n \frac{1}{\prod_{k=1}^p h_{[i]k}} Y_{i+p+1} K \left(\frac{y_1 - Y_{i+1}}{h_{[i]1}}, \dots, \frac{y_p - Y_{i+p}}{h_{[i]p}} \right) \left[\sum_{i=1}^n \frac{1}{\prod_{k=1}^p h_{[i]k}} K \left(\frac{y_1 - Y_{i+1}}{h_{[i]1}}, \dots, \frac{y_p - Y_{i+p}}{h_{[i]p}} \right) \right]^{-1}.$$

Пример 2.8. *Фильтрация сигнала с аддитивным шумом.* Пусть наблюдается дискретный сигнал $Y(t)$ с аддитивным шумом. Представим сигнал в виде: $Y(t) = Y(t-1) + \Delta Y(t)$, где $\Delta Y(t) = (Y(t) - Y(t-1)) + \xi(t)$ (зашумленное изменение сигнала на интервале $[t-1, t]$) и введем обозначения: $t_i = \overline{t(T-n-p+i)}$ - моменты времени, в которые производились измерения сигнала на интервале наблюдения длительностью T ; $\Delta Y^i = \Delta Y(t(T-n-p+i))$, $i = \overline{1, n+p}$ - выборочные значения изменения сигнала, величина p определяет порядок авторегрессионной зависимости, т.е. определяет степень влияния прошлых реализаций процесса $\Delta Y(t)$ на его значение в настоящий момент.

Рассматривая изменения сигнала $\Delta Y(t(i))$ как функцию $F(\cdot, \cdot)$ от времени и от значений на предыдущих интервалах времени, т.е. $\Delta Y(t(i)) = F\{t(i), \Delta Y(t(i-1)), \dots, \Delta Y(t(i-p))\}$, и, применяя непараметрические оценки, полученные в [43, 104], оценка $\hat{Y}(t(T))$ значений сигнала $Y(t(T))$ в момент времени $t(T)$ строится по следующим данным: $\{t_1, \dots, t_{N+p}\}$, $\{\Delta Y_1, \dots, \Delta Y_{N+p}\}$, где $t_i = t(T-N-p+i)$, $i = \overline{1, N+p}$ - моменты времени, в которые производились измерения; $\Delta Y_i = \Delta Y(t(T-N-p+i))$, $i = \overline{1, N+p}$ - выборочные значения изменения сигнала. Оценка $\hat{F} = \hat{F}(\Delta Y_N, \dots, \Delta Y_{N+p-1}, t(T))$ функции $F(\cdot)$ в момент времени $t(T)$ имеет вид:

$$\hat{F} = \frac{\sum_{i=1}^N \frac{\Delta Y_{i+p}}{h_{[i]}^t \prod_{z=1}^p h_{[i]z}^\Delta} K \left(\frac{\Delta Y_N - \Delta Y_i}{h_{[i]1}^\Delta}, \frac{\Delta Y_{N+1} - \Delta Y_{i+1}}{h_{[i]2}^\Delta}, \dots, \frac{\Delta Y_{N+p-1} - \Delta Y_{i+p-1}}{h_{[i]p}^\Delta}, \frac{t(T) - t_{i+p}}{h_{[i]}^t} \right)}{\sum_{i=1}^N \frac{1}{h_{[i]}^t \prod_{z=1}^p h_{[i]z}^\Delta} K \left(\frac{\Delta Y_N - \Delta Y_i}{h_{[i]1}^\Delta}, \frac{\Delta Y_{N+1} - \Delta Y_{i+1}}{h_{[i]2}^\Delta}, \dots, \frac{\Delta Y_{N+p-1} - \Delta Y_{i+p-1}}{h_{[i]p}^\Delta}, \frac{t(T) - t_{i+p}}{h_{[i]}^t} \right)}. \quad (2.24)$$

Функция регрессии обладает фильтрующим свойством, поскольку она строит условное математическое ожидание, которое является усредненным по

наблюдениям значением сигнала, то есть $\hat{F}(\Delta Y_N, \dots, \Delta Y_{N+p-1}, t(T))$ является отфильтрованным значением изменения сигнала в момент времени $t(T)$.

Для построения прогноза величины $\Delta \hat{Y}$ на интервале $[t(T), t(T+1)]$ необходимо вычислить функцию $\hat{F}(\cdot)$ в точке $(\Delta Y_{N+1}, \dots, \Delta Y_{N+p}, t(T+1))$. Таким образом, прогноз изменения сигнала будет выглядеть так:

$$\Delta \hat{Y}(t(T+1)) = \hat{F}(\Delta Y_{N+1}, \dots, \Delta Y_{N+p}, t(T+1)). \quad (2.25)$$

2.6. Метод цифрового дифференцирования зашумленного временного ряда на основе модели состояний ДО

Задача цифрового дифференцирования измеряемых в реальном масштабе времени зашумленных рядов с неизвестными характеристиками встречается при решении многих прикладных задач, и, несмотря на большое количество работ, общего формализованного подхода не существует. В традиционной постановке задачи дифференцирования предполагается выполнение условий¹⁸:

1) сигнал S является некоторой непрерывной и хотя бы один раз дифференцируемой функцией s времени t , удовлетворяющей равенству $S=s(t)$;

2) измерения сигнала S осуществляются в дискретные и равноотстоящие моменты времени t_k , $k=1,2,\dots$ связанные соотношением вида $t_k=t_k+\Delta t$, где Δt – некоторый постоянный интервал времени t , $t_0=0$;

3) измеренные значения $Y_k = Y(t_k)$ сигнала S состоят из истинных значений $s_k = s(t_k)$ сигнала S и ошибок его измерения $\varepsilon_k = \varepsilon(t_k)$: $Y_k = S_k + \varepsilon_k$

4) ошибки измерения образуют последовательность независимых и одинаково распределенных случайных величин, удовлетворяющих условиям вида: $M(\varepsilon_k) = 0$, $M(\varepsilon_k^2) = \sigma_k^2$, $M(\varepsilon_k \varepsilon_{k-1}) = 0$.

5) в каждый момент времени t_k у нас имеются p измеренных значений $Y_{k-p+1}, Y_{k-p+2}, \dots, Y_{k-1}, Y_k$.

Требуется по имеющимся в момент времени t_k измеренным значениям $Y_k = Y(t_k)$ вычислять оценку \hat{P}_k его производной P_k , определяемой равенством

$P_k = \frac{dS(t_k)}{dt}$, таким образом, чтобы близость $\hat{P}_k \approx P_k$ обеспечивалась в смысле

выбранного показателя качества (критерия).

¹⁸ А.В. Майстренко, А.А. Светлаков. Методы и алгоритмы цифрового дифференцирования сигналов в системах реального времени. - Томск : ТУСУР, 2009. – 138 с.

Далее изложим два новых подхода к получению значения оценки производной и сравним их, используя следующие показатели качества вычисления производной.

Коэффициент детерминации - это квадрат множественного коэффициента корреляции (одна из форм):

$$R^2 = \frac{\sum_{k=1}^N (\hat{y}_k - M[y_k])^2}{\sum_{k=1}^N (y_k - M[y_k])^2}, \left(R'^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \right), \quad (2.26)$$

где \hat{y}_k - модельные значения; y_k - уровни временного ряда; $M[\cdot]$ - оператор математического ожидания; $\bar{y}_i = \frac{1}{N} \sum_{i=1}^N y_i$, N - объем выборки; $R \in [0; 1]$.

Среднеквадратическая относительная ошибка (СКО):

$$\gamma_1 = \gamma_{CKO} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{\max(y_i)} \right|^2}, \hat{y}_i = f(x_i). \quad (2.27)$$

Среднее относительное отклонение (СОО):

$$\gamma_2 = \gamma_{COO} = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{\max(y_i)}. \quad (2.28)$$

Максимальное относительное отклонение (МОО):

$$\gamma_3 = \gamma_{MOO} = \max_{i=1, N} \frac{|y_i - \hat{y}_i|}{\max(y_i)}. \quad (2.29)$$

В качестве показателя уровня шума используется относительная погрешность γ , вычисленная по следующей формуле:

$$\gamma = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%, \quad (2.30)$$

где \hat{y}_i - значение сигнала, уровень шума в котором измеряется; y_i - истинное значение сигнала; N - число точек.

Функция регрессии (п. 2.8) обладает свойством фильтрации, поскольку она строит условное математическое ожидание, которое является усредненным по наблюдениям значением сигнала, то величина $\hat{F}(\Delta Y_N, \dots, \Delta Y_{N+p-1}, t(T))$ - является отфильтрованным значением изменения сигнала в момент времени $t(T)$.

Следовательно, данный подход можно использовать при решении задачи цифрового дифференцирования сигналов.

Для получения прогнозного значения оценки производной сигнала на один шаг после момента $t(T)$ необходимо вычислить значения оценки функции $F(\cdot)$ в точке $t(T+1)$. Оценка производной на момент $t(T+1)$, согласно (2.24), имеет вид:

$$\hat{F}(t(T+1)) = \frac{\sum_{i=1}^N \frac{\Delta Y_{i+p}}{h_{[i]}^t \prod_{z=1}^p h_{[i]}^{\Delta z}} K\left(\frac{t(T+1)-t_{i+p}}{h_{[i]}^t}\right)}{\sum_{i=1}^N \frac{1}{h_{[i]}^t \prod_{z=1}^p h_{[i]}^{\Delta z}} K\left(\frac{t(T+1)-t_{i+p}}{h_{[i]}^t}\right)}. \quad (2.31)$$

На рис. 2.7 приведена структурная схема процедуры оценивания производной и алгоритм оценивания производной, соответственно.

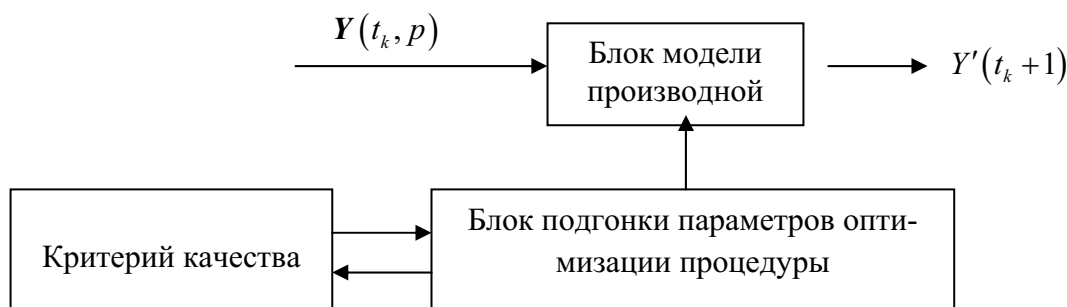


Рис. 2.7. – Структурная схема построения оценки производной сигнала. Обозначения: $Y(t_k, p) = (Y_{k-p+1}, Y_{k-p+2}, \dots, Y_{k-1}, Y_k)$, $Y'(t+1)$ - прогнозируемое значение величины Y в момент времени $t+1$

На рис. 2.7 блок подгонки параметров оптимизации процедуры реализует процедуру выбора параметров (размытости) в смысле оптимума функционала качества, сформированного в блоке задания критерия качества. В блоке построения модели сигнала осуществляется вычисление оценки регрессии с использованием оптимальных параметров.

Алгоритм 2.9. Оценка производной на основе непараметрического подхода

Вход: наблюдения дискретного сигнала $Y(t)$ глубины p :
 $Y(t_j, p) = (Y_{j-p+1}, Y_{j-p+2}, \dots, Y_{j-1}, Y_j)$, $j = 1, 2, \dots$; $t_j = t(T - N - p + j)$, $j = \overline{1, N + p}$ - моменты времени, в которые производились измерения.

Выход: оценка изменения сигнала $\Delta \hat{Y}$ на интервале $[t(T), t(T+1)]$.

1: Проверка качества параметров (размытости) по критерию ММП [43].

2: Вычисление $\hat{F}(\Delta i_{s\alpha}^N, \dots, \Delta i_{s\alpha}^{N+p-1}, t(T))$ по формуле (2.24).

3: Вычисление $\Delta \hat{Y} = \hat{Y}(t(T)) - Y(t(T-1)) = \hat{F}(\Delta Y^N, \dots, \Delta Y^{N+p-1}, t(T))$

Вернуть $\Delta \hat{Y}$.

Метод цифрового дифференцирования измеряемых координат на основе параметризации авторегрессии ARADS. Алгоритм, реализующий метод, основан на установлении соответствия между значениями коэффициентов разностной схемы, однозначно соответствующих заданным кривым (область допустимых значений), и значениями коэффициентов AR-моделей (рис. 2.5, 2.8).

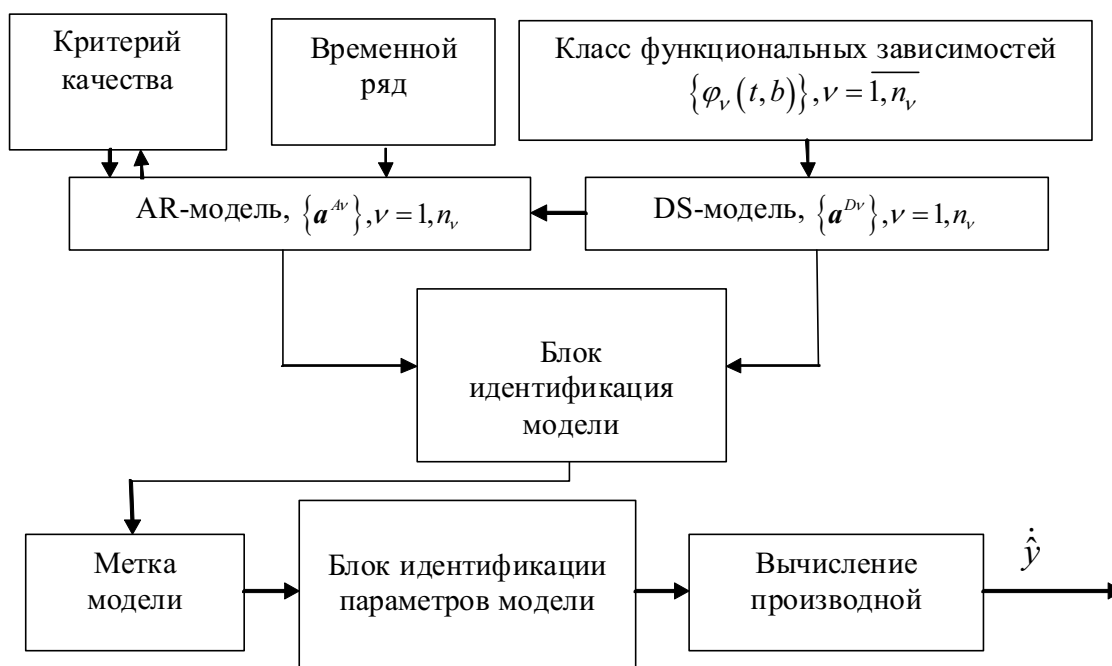


Рис. 2.8. – Структурная схема построения оценки производной сигнала на основе метода ARADS

В качестве аппроксимирующей наблюдения $y = y(t)$ функции $\hat{y} = \hat{y}(t)$ используется модель функциональной зависимости $f(t, b)$ (b - вектор параметров выбранной модели), определяемая из условия минимального евклидова расстояния между указанными коэффициентами AR-моделей функциональных зависимостей (кривых), полученных по наблюдаемым значениям ВР и областью допустимых значений (коэффициентов разностной схемы). Значение производной $\dot{f}_{v_0}(t, b)$ выбранной модели полагается равной производной функции $\dot{y} = \hat{y} = \dot{f}_{v_0}(t, b)$.

Алгоритм 2.10. Адаптивная идентификация модели и дифференцирование на основе ARADS

Вход: множество наблюдаемых пар: (t_j, \bar{y}_j) , $y_j = Y(t_j)$, $j = \overline{1, k}$, где \bar{y}_j – усредненное взвешенное значение глубины k : $\bar{y}_j = \frac{1}{k} \sum_{e=t_j-k+1}^{t_j} \omega_e y_e$; ε – точность

подгонки параметров функциональной модели; признаки состояний ДО по алгоритму 2.6 (ОДЗ $\mathbf{a}^{Dv} = (a_1^{Dv}, \dots, a_{p^v}^{Dv})$ для заданного множества кривых $f_v, v = \overline{1, n_f}$).

Выход: значение оценки производной $\dot{y}(t)$ в точке t .

1: Строим n_f AR-моделей (для каждой кривой $f_v, v = \overline{1, n_f}$) по выборке (t_j, \bar{y}_j) , определяем коэффициенты AR-моделей $(a_1^{Av}, \dots, a_{p^v}^{Av}), v = \overline{1, n_f}$:

$$(a_1^{Av}, \dots, a_{p^v}^{Av}) = \arg \min_{a_j^{Av} \in R} \sum_{l \in \{1, \dots, k\}} \phi \left(y_l - \sum_{j=1}^{p^v} a_j^{Av} y_{l-j} \right).$$

где ϕ вышеопределенная функция (утверждение 2.3).

2: Находим расстояние $r(\mathbf{a}^{Dv}, \mathbf{a}^{Av})$ между векторами $(a_1^{Dv}, \dots, a_{p^v}^{Dv}), (a_1^{Av}, \dots, a_{p^v}^{Av})$: $r(\mathbf{a}^{Dv}, \mathbf{a}^{Av}) = \sqrt{\sum_{j=1}^{p^v} (a_j^{Dv} - a_j^{Av})^2}$ в p^v -мерном базисном пространстве, координатами которого служат коэффициенты разностной схемы.

3: Принимаем решение в пользу v_0 -й модели $f_{v_0}(t, \mathbf{b})$ функциональной зависимости $v_0 = \arg \min_v r(\mathbf{a}^{Dv}, \mathbf{a}^{Av})$.

4: Осуществляем подгонку параметров выбранной модели, решая условную систему алгебраических неравенств: $|f_{v_0}(t_j, \mathbf{b}) - y(t_j)| < \varepsilon, j = \overline{1, k}$.

Вернуть $\dot{y} = \dot{f}_{v_0}(t, \mathbf{b})$ в точке t .

В табл. 2.6 и на рис. 2.9-2.11 представлены результаты сравнительного моделирования изложенных нетрадиционных методов оценивания производной сигнала с некоторыми используемыми на практике (на основе скользящего среднего, на основе ряда Тейлора и полиномиальной интерполяции).

Таблица 2.6 – Качество оценивания производной по временному ряду

Шум	Метод	СКО	СОО	МОО	R^2
5%	Аппроксимация на основе	0.398	0.158	1.007	0.981

	ряда Тейлора				
	Скользящее среднее	0.394	0.155	1.099	0.970
	Интерполяция полиномом	0.412	0.169	1.771	0.895
	Ядерная фильтрация	0.453	0.205	1.084	0.986
	ARADS-метод	0.194	0.153	1.271	0.996
27%	Аппроксимация на основе ряда Тейлора	0.899	0.809	3.25	0.574
	Скользящее среднее	0.858	0.736	2.258	0.596
	Интерполяция полиномом	0.414	0.172	1.745	0.899
	Ядерная фильтрация	0.697	0.485	2.026	0.768
	ARADS-метод	0.313	0.158	1.343	0.975
55%	Аппроксимация на основе ряда Тейлора	1.228	1.513	4.859	0.354
	Скользящее среднее	1.144	1.309	4.631	0.381
	Интерполяция полиномом	0.414	0.171	1.768	0.893
	Ядерная фильтрация	0.686	0.470	1.593	0.885
	ARADS-метод	0.314	0.171	1.368	0.989
75%	Аппроксимация на основе ряда Тейлора	1.417	2.009	7.719	0.266
	Скользящее среднее	1.327	1.759	7.155	0.289
	Интерполяция полиномом	0.417	0.174	1.626	0.936
	Ядерная фильтрация	0.739	0.547	2.281	0.897
	ARADS-метод	0.315	0.164	1.452	0.997

Замечание. Для зашумленного временного ряда целесообразно предварительное сглаживание (например, скользящим усреднением). Данные в табл. 2.6 приведены без предварительной подготовки временного ряда.

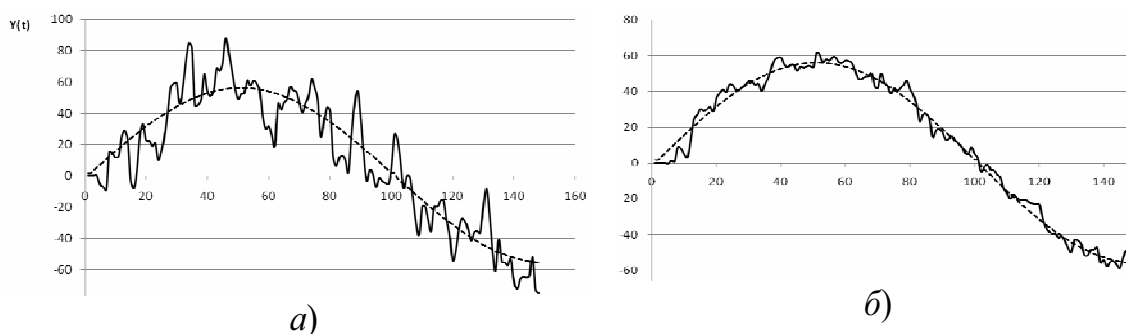


Рис. 2.9. а) Временной ряд (дискретизация модельной функции - пунктирная линия) с аддитивным 27%-м шумом (сплошная линия), б) прогнозное значение ряда с 19% ошибкой, по которому вычисляется производная

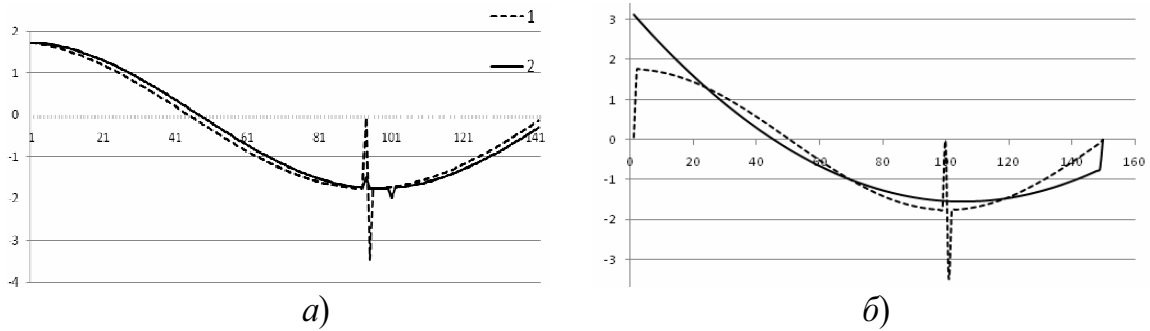


Рис. 2.10. а) Производная (сплошная линия), найденная по методу ARADS (2); б) производная на основе на основе полиномиальной интерполяции

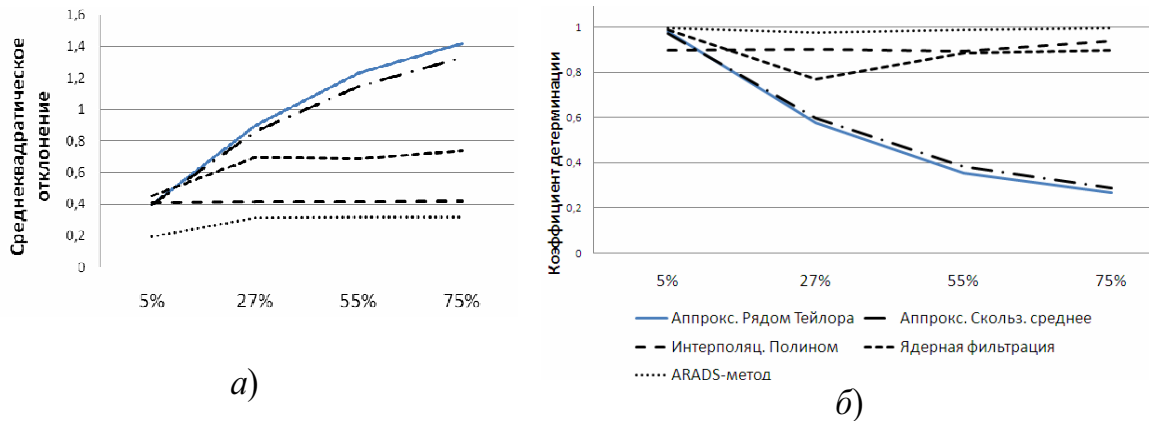


Рис. 2.11. а) СКО значения производной при различных значениях уровня шума по формуле (2.30); б) коэффициент детерминации по формуле (2.26)

Для случая нормально распределенного шума статистическая значимость указанных данных удовлетворяет t -тесту на уровне 0.05.

2.7. Определение типа аттрактора и времени вхождения в аттрактор

В синергетическом подходе при построении адаптивных регуляторов (гл. 7) длительность времени вхождения в аттрактор является параметром функционала качества и на практике определяется эмпирически. Ниже на основе результатов [121] и алгоритмах 2.1, 2.5 предлагается алгоритм 2.11 для определения типа аттрактора и времени вхождения в аттрактор.

Обозначим совокупность наблюдаемых ВР через $Y=(y_1, \dots, y_n)$, $y_j=(y_1^j, \dots, y_N^j)$, где n – порядок системы ОДУ, y_j – j -я координата ДО, N – размер ВР. Сформируем одномерный ВР, состоящий из разностей соседних значений исходного ВР Y [5, 121, 166]:

$$Z_1=(z_1, \dots, z_{N-1}), z_l=\sqrt{\sum_{j=1}^n (y_l^j - y_{l+1}^j)^2}, l=\overline{1, N}. \quad (2.32)$$

Алгоритм 2.11. Определение типа аттрактора и времени вхождения в аттрактор

Вход: обучающая выборка в виде временных рядов, элементы которых состоят из (реконструированных) решений исходной системы ОДУ – модели ДО.

Выход: t^* - момент вхождения в аттрактор и тип аттрактора, вид функциональной зависимости равновесного состояния.

- 1: Формируем временной ряд (2.32).
- 2: Строим AR-модель для ВР (2.32) на основе алгоритма 2.5.
- 3: Определяем AR-коэффициенты $a(t)$ и корни характеристического полинома $q(t)$ AR-модели ряда (2.32), а также среднее и дисперсию ряда (2.32).
- 4: Формируем решающие правила согласно следующему списку условий:

Состояние равновесия: узел, фокус, седловая точка.	<ol style="list-style-type: none"> 1) функции $a(t)$ и $q(t)$ являются убывающими, начиная с некоторого момента времени t; 2) $Z_1(t)$ убывает (убывает, осциллируя с затухающей амплитудой); $\lim_{t \rightarrow \infty} Z_1 = 0$.
--	---

Состояние равновесия: замкнутый предельный цикл.	<ol style="list-style-type: none"> 1) $a(t) \rightarrow \text{const}$, $\max q(t) \rightarrow 1$; 2) Z_1 – периодический ряд;
--	--

Состояние равновесия не существует.	<ol style="list-style-type: none"> 1) $a(t)$ хаотично изменяется; максимальный по модулю корень характеристического полинома $q_0(t)$ может иметь разный характер ($q_0(t) > 1$, $q_0(t) < 1$, $q_0(t) = 1$); 2) $\lim_{t \rightarrow \infty} Z_1$ не существует.
-------------------------------------	---

5: Ищем момент t^* по правилу: $t^* = \begin{cases} \arg \min_l (|Z_1| < \varepsilon), \exists l \in [1, N-1]: |Z_1| < \varepsilon, \\ 0, & -\exists l \in [1, N-1]: |Z_1| < \varepsilon. \end{cases}$

6: Определяем уровень статистической достоверности полученных результатов.

7: Идентифицируем вид функциональной зависимости $Z_1(t)$ для $t > t^*$ по алгоритму 2.6.

Вернуть: решающие правила (п. 4); t^* - момент вхождения в аттрактор (если он существует); функциональную модель временного ряда (2.32) для оценки скорости движения сложного ДО к аттрактору.

Результаты алгоритма 2.11 используются при нахождении параметров регулятора управления сложным ДО на основе метода АКАР в гл. 7, 9.

2.8. Численное моделирование распознавания состояний ДО и оценка его качества

Численное моделирование проведено на примерах нескольких временных рядов, являющимися моделями некоторых характеристик технических систем по имитационной модели, разработанной в среде Matlab.

Экспериментальное исследование качества сглаживания временного ряда проводилось в сравнении с результатами непараметрической идентификации нелинейной регрессионной модели сигнала (пп. 2.6, 2.7). В [86] были проведено численное моделирование алгоритмов фильтрации, прогнозирования и распознавания состояний измеряемых сигналов ДО по методам гл. 2 (рис. 2.12).

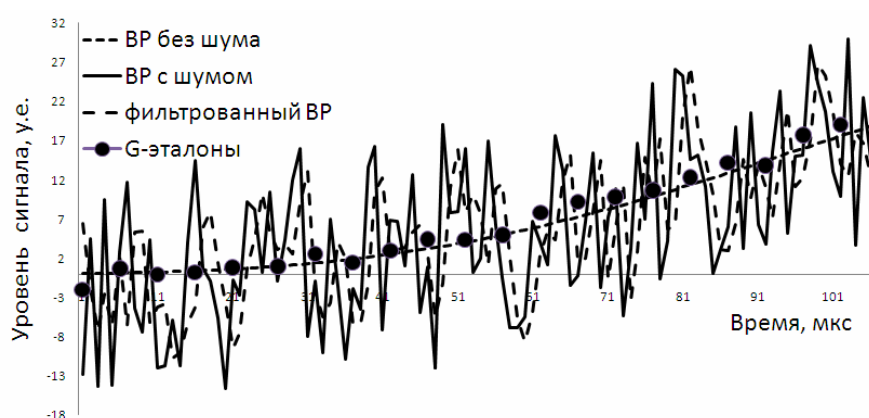


Рис.2.12. Результат фильтрации модельного сигнала; по оси OX – дискретные отсчеты времени, OY – уровень сигнала, BP – временной ряд

На модельный сигнал накладывались различные шумы уровня, не превышающего 30 % от максимального значения полезного сигнала: равномерный шум на интервалах $[-5;5]$, $[-10;10]$, $[-150;150]$, $[-300;300]$; гауссовский шум с параметрами $N(0,10)$, $N(0,50)$, $N(0,10000)$, $N(0,50000)$. По полученным данным (временным рядам) были построены фильтрованные значения по 1000 точкам. В качестве показателя уровня шума использована относительная погрешность γ ,

вычисленная по формуле:
$$\gamma = \frac{1}{1000} \sum_{i=1}^{1000} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \%,$$
 где \hat{y}_i – значение сигнала,

уровень шума в котором измеряется; y_i – истинное значение сигнала. На рисунке 2.12 показана удовлетворительная точность сглаживания временного ряда, тогда как непараметрическая ядерная фильтрация дает устойчивый эффект не выше 14-ти процентов против 30%-й первоначальной зашумленности по 5%-му уровню значимости.

Сравнение методов распознавания состояний ДО по точности и времени с применением предварительной фильтрации показало, что время распознавания с фильтрацией больше на 10^1 - 10^2 мкс при данных для окна в 50 отсчетов и ряда из

150 отсчетов; точность распознавания при использовании стандартной процедуры квантования (п.1.2, опр. 5.2) для данных нелинейных моделей без фильтрации выше, чем точность распознавания для данных с фильтрацией, примерно на 1-3%; точность распознавания для адаптивной процедуры (гл. 5) квантования при использовании фильтрации для 10%-го шума в среднем выше на 3-10 %, чем без фильтрации. Вышесказанное убеждает, что необходимы дополнительные исследования к выяснению условий применения / неприменения предварительной фильтрации к временным рядам с нелинейным трендом.

При сравнении эффективности распознавания в зависимости от вида тренда временного ряда (при размахе значений элементов рассмотренных рядов более 2400 у.ед. и 5%-м шуме) для «немонотонных» моделей результаты распознавания близки к «ожидаемым» (~100%).

Результаты дополнительных экспериментальных исследований изложены в гл. 5 при анализе разных типов квантования исходной траектории процесса.

Выводы. В разделе показано применение эффективного нетрадиционного метода формирования признаков состояний ДО по сопровождающему его временному ряду, фильтрующие свойства которых позволяют выделить характерные черты состояний динамического объекта.

На основе обобщенных эталонов состояний возможно решение трех задач: распознавание состояний ДО, модель характеристик которого задана в виде многомерного временного ряда; сглаживание временного ряда; идентификация тренда фрагмента временного ряда, соотнесенного с состоянием ДО (признака формы).

Показано, что идентификацию состояний динамического объекта возможно осуществлять по значениям коэффициентов авторегрессионной модели, построенной по обобщенным эталонам. Результаты экспериментальных исследований утверждают, что изложенные алгоритмы позволяют уверенно распознавать состояния динамических объектов с неопределенностью.

Результаты, представленные во 2-й главе, опубликованы в работах [75, 80-89, 90].

«Мы находимся в положении, несколько аналогичном положению человека, держащего связку ключей и пытающегося открыть одну за другой несколько дверей. Рано или поздно ему удастся подобрать ключ к очередной двери, но сомнения относительно взаимно-однозначного соответствия между ключами и дверями у него остаются» /Ю. Вигнер¹⁹/

Глава 3. МОДИФИКАЦИЯ МЕТОДА АНАЛИЗА ИЕРАРХИЙ ДЛЯ ДИНАМИЧЕСКИХ НАБОРОВ АЛЬТЕРНАТИВ

Задачи «наилучшего» выбора относятся к классу слабоформализованных (некорректных), поскольку отсутствует общий теоретически обоснованный подход на основе достоверного критерия качества выбираемого решения. Из программы исследования некорректных (эвристических) процедур основателя алгебраической теории алгоритмов [54-56] следует, что «задача состоит в том, чтобы, признав как реальность их практическую пользу для решения плохо формализованных проблем, изучить с помощью строгих математических методов само множество таких процедур».

В силу важности метода ММАИ (рис. 3.1) среди известных («Электра» и др.) для решения слабоформализованных задач на основе прикладных интеллектуальных систем, изложим его теоретические предпосылки и иллюстрируем примерами.

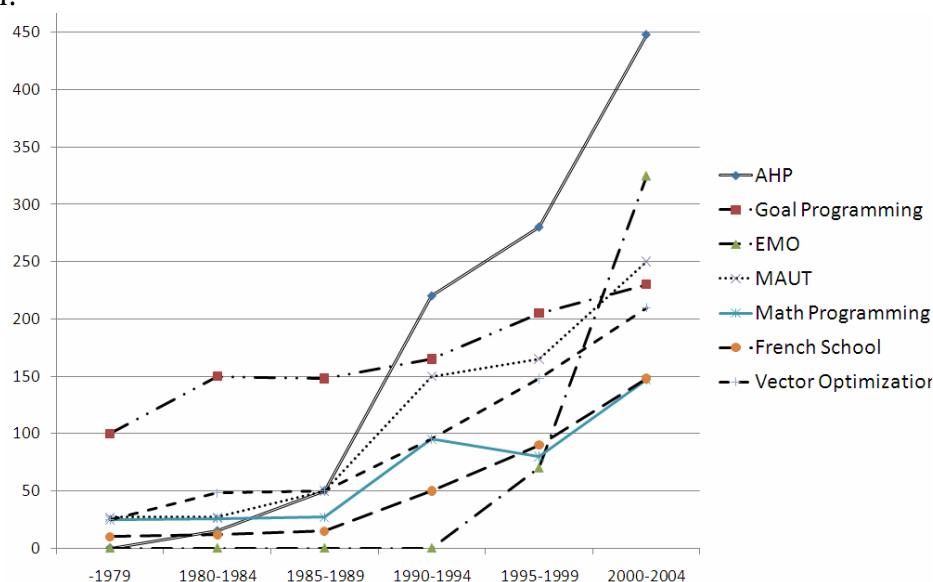


Рис.3.1. История числа публикаций по разным многокритериальным методам по данным базы данных ISI²⁰ (AHP - метод аналитических иерархий)

¹⁹ Лауреат Нобелевской премии по физике в 1963г.

²⁰ Реферативно-библиографическая база данных Института научной информации США <http://isiwebofknowledge.com/>

На рис. 3.1 используются обозначения: метод аналитических иерархий (АИР), целевое программирование (Goal Programming), эволюционная многокритериальная оптимизация (ЕМО), многокритериальная теория полезности (МАУТ), метод математического программирования (Math Programming), различные человеко-машинные процедуры (в частности Vector Optimization).

3.1. История вопроса. Нежелательные эффекты МАИ

Метод анализа иерархий (МАИ) является одним из самых популярных методов, используемых в информационно-аналитических и интеллектуальных системах поддержки принятия решений, позволяющим структурировать проблему, оценивать построенный набор альтернатив по каждому из заданных критериев, находить неточности и противоречия в суждениях эксперта, ранжировать альтернативы [141]. По ориентировочным оценкам число статей прикладного характера, в которых метод анализа иерархий применяется к решению прикладных слабоструктурированных многокритериальных задач, составляет более полутора тысяч. Этому способствует и пакет зарубежных программ EXPERT CHOICE, реализующий МАИ.

Как известно, недостатком «классического» МАИ, отмеченным и самим автором метода, являются два противоречия, первое из которых связано с эффектом единичной нормировки, приводящей к тому, что предпочтения, выявленные на всем множестве альтернатив (критериев), могут не совпадать с «частными» предпочтениями на подмножестве альтернатив (критериев). Содержательно это означает следующее. При оценивании двух проектов z_1 и z_2 несколькими экспертами по методу МАИ устанавливается, что z_1 предпочтительней z_2 ($z_1 \succ z_2$), знак « \succ » означает факт предпочтительности (доминирования). При появлении третьего проекта система (на основе МАИ) их ранжирует по ценности заново, и в результате возможна ситуация: z_2 предпочтительней z_1 ($z_2 \succ z_1$).

Отметим, что согласно аксиоме К. Эрроу²¹ при неизменных критериях оценки отношение предпочтения любых двух элементов множества альтернатив постоянно, т.е. бинарное отношение предпочтения альтернатив не зависит от дополнительных альтернатив.

Второе противоречие связано с применением в МАИ линейной свертки критериев. В работе [126] отмечено, что в определенном классе многокритериальных задач при помощи линейной свертки в принципе невозможно найти «наилучшее» решение, и поэтому ее использование становится некорректным. Линейная свертка основана на неявном принципе: «низкая оценка по одному критерию может быть компенсирована высокой оценкой по другому критерию».

²¹ К. Дж. Эрроу получил в 1972 г. Нобелевскую премию за доказательство невозможности существования идеальной системы выборов.

Можно привести содержательные примеры (ниже примеры 3.1-3.3), когда этот принцип далеко неверен и даже абсурден (ухудшение качества изображения фильма не может быть компенсировано улучшением качества его звука).

При выборе наилучшего способа решения задачи многокритериального выбора следует руководствоваться принципом [132, 133]: всякое решение, которое потенциально может оказаться выбранным в результате решения многокритериальной задачи, может быть получено в результате максимизации используемой свертки критериев с некоторыми неотрицательными коэффициентами c_i , $\sum_i c_i = 1$. Другими словами, у каждого элемента, принадлежащего оптимальной (по Парето) области множества X , должна существовать возможность быть выбранным в качестве наилучшего решения. В противном случае, «наилучшее» решение может и не быть найдено вообще в смысле экстремума выбранного способа скаляризации (см. ниже примеры 3.1-3.3), в частности, в результате максимизации линейной свертки критериев в МАИ. Причиной этого может быть их нелинейный характер: результаты действия двух или нескольких факторов не складываются, а сложным образом (неизвестным) комбинируются (результат одновременного приема двух «хороших» по отдельности лекарств может быть отрицательным).

Таким образом, первая проблема МАИ связана с неточностью оценивания динамических наборов сравниваемых альтернатив при их пополнении (убавлении), вторая проблема связана с неточностью, вызванной применением линейной свертки при получении весовых коэффициентов альтернатив.

В продолжение исследований, проводившихся в работах [141-143], далее даются свойства модифицированной процедуры МАИ (ММАИ), а именно, доказано, что: 1) предпочтения, индуцированные на множествах (альтернатив, признаков) Θ_1 и Θ_2 ($\Theta_1 \subset \Theta_2$) посредством применения классической процедуры МАИ Саати, в общем случае не совпадают; 2) в модифицированной процедуре МАИ указанное противоречие устраняется; 3) «наилучшее» решение, выбираемое по ММАИ в смысле нелинейной свертки критериев, принадлежит парето-оптимальному множеству.

3.2. Основные определения и понятия. Постановка задачи

В общем виде постановка задачи, решаемой МАИ [141], включает цель, альтернативы и критерии оценки альтернатив; требуется выбрать наилучшую альтернативу. После построения иерархической структуры (цели-критерии-альтернативы) система парных сравнений элементов каждого уровня приводит к результату, который может быть представлен в виде обратно-симметричной матрицы A , элемент которой a_{ij} есть интенсивность проявления альтернативы (элемента иерархии) i относительно альтернативы j в смысле выбранного фиксиро-

ванного критерия. Главные собственные векторы матрицы \mathbf{A} парных сравнений интерпретируются как векторы приоритетов сравниваемых элементов. Для элементов каждого уровня вычисляются коэффициенты важности для каждой из альтернатив в виде произведения в случае, когда оценки элементов заданы в виде отношения (мультипликативный МАИ).

Обозначим формируемую на каждом этапе матрицу парных сравнений (МПС) $\mathbf{A} = \|a_{ij}\|_{g \times g}$ альтернатив (далее в контексте работы будем предполагать понятия альтернативы и признака идентичными) $\Theta = (z_1, z_2, \dots, z_g)$. Матрица $\mathbf{A} = \|a_{ij}\|_{g \times g}$, $a_{ij} = w_i w_j^{-1}$ относительных весовых коэффициентов w_i, w_j - компонент весового вектора $\mathbf{w} = (w_1, w_2, \dots, w_g)^T$, где g - количество сравниваемых альтернатив должна обладать свойствами [141]:

$$a) a_{ij} > 0; b) a_{ij} = a_{ji}^{-1}; c) a_{ij} = a_{ik} \cdot a_{kj}, i, j = \overline{1, g}; d) \mathbf{A} \cdot \mathbf{w} = g \cdot \mathbf{w}, \quad (3.1)$$

число g является максимальным собственным значением матрицы \mathbf{A} , и для некоторого единственного (нормированного) вектор-столбца $\mathbf{w} = (w_1, \dots, w_g)^T$ с положительными компонентами выполняется равенство $d)$, при этом условие $c)$ - условие совместности (согласованности) МПС \mathbf{A} .

Замечание. Известно [126, 141], что весовой вектор \mathbf{w} является собственным вектором совместной матрицы \mathbf{A} , отвечающим ее максимальному собственному значению n . Другими словами, в случае совместной матрицы весовой вектор необходимо является указанным собственным вектором. Однако, несмотря на то, что на практике требование совместности МПС часто не соблюдается, весовой вектор по-прежнему предлагается находить как собственный вектор уже несовместной МПС. В качестве обоснования такого решения предлагается специальный числовой показатель «индекс совместности», характеризующий степень доверия к полученным с помощью МАИ результатам. Индекс совместности трактуется как мера отклонения исходной несовместной матрицы от некоторой совместной [141]: при достаточно малом значении индекса совместности матрица парных сравнений «близка» к некоторой (совместной) матрице с нулевым значением этого индекса. Если индекс совместности превышает «пороговое» значение, то сделать вывод о близости указанных матриц нельзя, поэтому и применять МАИ в таких случаях не рекомендуется. В [126] отмечено, что по значению индекса совместности можно лишь опосредованно судить о величине итоговой «модельной» ошибки; точно она никогда и никем не может быть определена.

Обозначим $X = \{x\}$ - множество выбираемых решений - альтернатив $x, f(x)$ - числовой критерий на множестве X .

Решение $x' \in X$ называется оптимальным по Парето (парето-оптимальным) [133], если не существует такого возможного решения $x \in X$, для которого имеет место неравенство $f(x) > f(x')$. Все парето-оптимальные решения образуют множество Парето, обозначаемое $P_f(X)$.

Постановка задачи. Требуется построить процедуру оценки значимости альтернатив по совокупности заданных критериев, во-первых, сохраняющей бинарные отношения (предпочтения) при изменении (добавлении дополнительной альтернативы, или удалении альтернативы из исследуемой совокупности); во-вторых, гарантирующей корректное применение используемой (линейной, в частности) свертки, а также определить ее свойства и класс решаемых задач.

Прежде чем перейти к решению задачи, обсудим примеры из работ [52, 125, 141], иллюстрирующие негативные стороны классического применения МАИ, приводящих к неточности в принятии решения.

Пример 3.1. Пусть задача состоит в выборе прямоугольного участка из следующих трех вариантов: (I) 7×15 , (II) 10×10 , (III) 5×20 (измерение производится в некоторых единицах длины). Первый участок, очевидно, имеет максимальную площадь. Однако, при использовании МАИ с равновесными метрическими критериями (длина и ширина участка) и аддитивной сверткой критериев, ошибочный выбор будет сделан в пользу второго участка. Действительно, соответствующие МПС имеют вид:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 7/10 & 7/5 \\ 10/7 & 1 & 10/5 \\ 5/7 & 5/10 & 1 \end{bmatrix}; \quad \mathbf{A}_2 = \begin{bmatrix} 1 & 15/10 & 15/20 \\ 10/15 & 1 & 10/20 \\ 20/15 & 20/10 & 1 \end{bmatrix}.$$

оценки собственных векторов МПС \mathbf{A}_1 и \mathbf{A}_2 равны: $w^1 = (0.318, 0.455, 0.227)$, $w^2 = (0.333, 0.222, 0.444)$, соответственно, итоговый весовой вектор равен: $w = (0.326, 0.338, 0.336)$.

Заметим, что при использовании мультипликативной свертки весовой вектор по равновесным критериям в примере 2 будет равен: $(0.318^{0.5} \cdot 0.333^{0.5}, 0.455^{0.5} \cdot 0.222^{0.5}, 0.227^{0.5} \cdot 0.444^{0.5}) = (0.326, 0.319, 0.319)$, или в нормализованном виде: $w = (0.338, 0.331, 0.331)$, что тоже соответствует верному выбору. Известно [126, 133], что максимальное значение мультипликативной свертки на некотором решении (в данном случае, точки с координатами (ширина, длина)) является достаточным условием парето-оптимальности этого решения.

Пример 3.2. Пусть оцениваются две альтернативы (признака): z_1, z_2 по двум равновесным мерам относительной важности со следующими МПС [141]:

$\mathbf{A}_1 = \begin{bmatrix} 1 & 1/2 \\ 2 & 1 \end{bmatrix}$; $\mathbf{A}_2 = \begin{bmatrix} 1 & 4 \\ 1/4 & 1 \end{bmatrix}$. В результате применения МАИ получим для альтернатив z_1, z_2 вектор ВКА $\mathbf{w}=(w_1, w_2)=(0.567, 0.433)$. Так как $w_1 > w_2$, то $z_1 \succ z_2$. При оценивании трех альтернатив z_1, z_2, z_3 по тем же двум мерам относительной важности со следующими МПС:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 1/2 & 3 \\ 2 & 1 & 6 \\ 1/3 & 1/6 & 1 \end{bmatrix}; \mathbf{A}_2 = \begin{bmatrix} 1 & 4 & 1/2 \\ 1/4 & 1 & 1/8 \\ 2 & 8 & 1 \end{bmatrix}.$$

В результате классического применения МАИ получаем: $\mathbf{w}=(w_1, w_2, w_3)=(0.304, 0.338, 0.358)$, т.е. $z_2 \succ z_1$.

Пример 3.3. Третий пример касается тестовых методов распознавания образов (основы одного из наиболее эффективных подходов к созданию интеллектуальных систем), использующих для принятия решений наборы (тесты) [55], содержащие меньшее количество признаков и с большим весом, где под весом теста понимается сумма весовых коэффициентов признаков (см. гл. 4). Новый метод к определению «весов» сравниваемых признаков, изложенный в работе [99], учитывает вклад признаков в распознающую способность теста с учетом их взаимозависимости и базируется на представлении совокупности всех различных пар объектов из разных классов (образов) для каждого признака в виде мультимножества, дальнейшем применении МАИ, использующего парные сравнения признаков на основе специальным образом выбранных мер их относительной важности. Ниже приведенный пример связан с целесообразностью корректного оценивания при изменении анализируемого набора признаков.

Пусть матрица \mathbf{Q} (объект-признак) содержит описание 6-ти объектов, в матрице \mathbf{R} указывается на соответствие номеров объектов и классов, которым объекты принадлежат. Для данного примера тупиковыми тестами (матрица \mathbf{T}) являются наборы признаков $\Theta_1=\{z_1, z_2, z_3\}$, $\Theta_2=\{z_1, z_2, z_4\}$ и $\Theta_3=\{z_2, z_3, z_4\}$.

$$\mathbf{Q} = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 & z_4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 2 \end{bmatrix} \end{matrix}; \mathbf{R} = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 & z_4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{bmatrix} \end{matrix}; \mathbf{T} = \begin{matrix} & \begin{matrix} z_1 & z_2 & z_3 & z_4 \end{matrix} \\ \begin{matrix} 1 \\ 1 \\ 0 \end{matrix} & \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \end{matrix}.$$

Рис. 3.2. Пример матриц «объект-признак» (\mathbf{Q}), «объект-класс» (\mathbf{R}), тестов (\mathbf{T})

Корректное применение тестового алгоритма по принципу совпадения значений всех признаков не позволит новый объект $\mathbf{O}=(0, 1, 2, 1)$ отнести к одному из выделенных классов, однако фрагмент $(0, 1)$, порождаемый набором $\Theta=\{z_1, z_2\}$, содержится в \mathbf{O} и соответствующих объектах из первого класса и не содержится в объектах из второго класса, что дает нам основание полагать, что распознаваемый объект более близок к первому классу. Непосредственное применение МАИ в методе определения весовых коэффициентов признаков может привести к ошибочному, а, следовательно, и к неточности в принятии решения.

В работах [142, 143] предлагается следующая схема решения проблемы единичной нормировки, согласно которой порядок предпочтения альтернатив на множестве метрических критериев строится на основе относительного (взаимного) превосходства.

Для каждой пары альтернатив с номерами (i, j) определяются нормированные по s -му критерию значения весовых коэффициентов альтернатив (ВКА) v_{is}, v_{js} , т.е. $v_{is}+v_{js}=1$. Далее вычисляются выражения $D_i = \sum_{s=1}^k c_s v_{is}$, $D_j = \sum_{s=1}^k c_s v_{js}$ и строится обратно-симметричная матрица из отношений $\gamma_{ij}=D_i/D_j$, а именно: $\mathbf{B}=\|b_{ij}\|$, $b_{ij}=\gamma_{ij}$, $b_{ji}=1/b_{ij}$. Величина γ_{ij} называется «коэффициентом относительного превосходства i -й альтернативы». Далее рекомендуется ненормированный вектор приоритетов альтернатив $\mathbf{w}=(w_1, \dots, w_g)^T$ находить как собственный вектор (как и в МАИ) из решения матричного уравнения: $\mathbf{B} \cdot \mathbf{W}=\lambda_{\max} \cdot \mathbf{W}$, где λ_{\max} - максимальное собственное значение матрицы \mathbf{B} .

Заметим, однако, что выполнение свойства транзитивности ($b_{ij}=b_{ik} \cdot b_{kj}$), являющееся весьма важным условием для данного вывода (см. утверждение), в данном случае не гарантируется. Действительно, обозначив через v'_{is}, v'_{js} ненормированные по s -му критерию значения ВКА, получаем выражения: $v_{is}=v'_{is}/(v'_{is}+v'_{js})$, $v_{js}=v'_{js}/(v'_{is}+v'_{js})$. Отсюда следует, величина γ_{ij} - это не просто коэффициент относительного превосходства i -й альтернативы, а по сравнению с j -й, т.е. $D_i=D_{i/j}, D_j=D_{j/i}$:

$$D_{i/j} = \sum_{s=1}^k c_s v_{is} = \sum_{s=1}^k c_s \frac{v'_{is}}{v'_{is} + v'_{js}}, \quad D_{j/i} = \sum_{s=1}^k c_s v_{js} = \sum_{s=1}^k c_s \frac{v'_{js}}{v'_{is} + v'_{js}}.$$

Тогда получаем: $b_{ik} \cdot b_{kj}=\gamma_{ik} \cdot \gamma_{kj}=(D_{i/k}/D_{k/i}) \cdot (D_{k/j}/D_{j/k})$; $b_{ij}=(D_{i/j}/D_{j/i})$ и вопрос о справедливости $b_{ij}=b_{ik} \cdot b_{kj}$, таким образом, остается открытым.

Утверждение 3.1. Вектор ВКА необходимо является собственным вектором, соответствующим максимальному собственному значению λ_{\max} только в случае совместной (согласованной) матрицы парных сравнений.

Доказательство утверждения (приложение 3) непосредственно следует из вида характеристического уравнения для \mathbf{B} , имеющего место при выполнении свойств матрицы парных сравнений (МПС) (3.1) [141]:

$$\det(\mathbf{B} - \lambda \mathbf{E}) = \det \begin{bmatrix} (1-\lambda) & 1 & \dots & 1 \\ 1 & (1-\lambda) & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & (1-\lambda) \end{bmatrix} = (-1)^g \lambda^{g-1} (\lambda - g) = 0,$$

из которого следует, что матрица \mathbf{B} будет иметь только два собственных значения (при выполнении условий (3.1)), при этом, значение g является максимальным. В противном случае собственный вектор несовместной МПС будет соответствовать максимальному собственному значению, которое строго больше g , а не равно g .

3.3. Нелинейная модификация МАИ и ее свойства

Укажем основные этапы ММАИ и перечислим ее свойства.

I. Построим МПС на каждом из ν этапов (по числу ν мер относительной важности одной альтернативы над другой). Результатом каждого s -го уровня иерархии является g – компонентный вектор нормализованных значений ВКА – $\mathbf{w}_s = (w_1^s, w_2^s, \dots, w_g^s)$. Введем весовые коэффициенты мер относительной важ-

ности альтернатив $c_s, s \in \{1, 2, \dots, \nu\}, \sum_{s=1}^k c_s = 1$.

II. Сформируем всевозможные векторы $\mathbf{w}_{ij}^s = (w_{ij}^s(i), w_{ij}^s(j))$ локальных ВКА уровня 1 по формулам 3-й строки табл. 1.

Таблица 3.1. Этапы модифицированного МАИ

	Промежуточные исходные данные s -го уровня иерархии, $(w_1^s, w_2^s, \dots, w_g^s)$	Относительные ВКА z_i, z_j	Вектор относительных нормализованных ВКА
I	$w_i^{ts} = \left(\prod_{l=1}^g a_{il}^s \right)^{1/g}, i = \overline{1, g}$	$w_i^s = w_i^{ts} / \sum_{l=1}^g w_l^{ts}, i = \overline{1, g}$	$(w_1^s, w_2^s, \dots, w_g^s),$ $s \in \{1, 2, \dots, \nu\}$
II	w_i^s, w_j^s	$w_{ij}^s(i) = \frac{w_i^s}{w_i^s + w_j^s}, w_{ij}^s(j) = \frac{w_j^s}{w_i^s + w_j^s}$	$\mathbf{w}_{ij}^s = (w_{ij}^s(i), w_{ij}^s(j))$

III	$w_{ij}(i) = \sum_{s=1}^k c_s w_{ij}^s(i), w_{ij}(j) = \sum_{s=1}^k c_s w_{ij}^s(j), \sum_{s=1}^k c_s = 1$	$(w_{ij}(i), w_{ij}(j))$	
IV	$u_i = \sum_{l=1}^g w_{il}(i)$	$V_i = u_i / \sum_{l=1}^g u_l, i = \overline{1, g}$	(V_1, V_2, \dots, V_g)

III. Сформируем матрицу $\mathbf{W} = \|\|w_{ij}\| = \|(w_{ij}(i), w_{ij}(j))\|$, где векторы $(w_{ij}(i), w_{ij}(j))$ – локальные нормализованные ВКА z_i, z_j уровня 2 относительно всей совокупности мер относительной важности признаков, компоненты которых находим по формулам 4-й строки табл.1.

IV. Глобальные значения ВКА считаем по одной из формул [141]:

$$V_i^{(1)} = \frac{1}{g} \sum_{j=1}^g w_{ij}(i), V_i^{(2)} = \left(\prod_{j=1}^g w_{ij}(i) \right)^{1/g}, i = \overline{1, g}.$$

Свойства обобщенной процедуры МАИ сформулируем в виде четырех теорем. Обозначим $\rho_1 = \{(z_i, z_j) : z_i \succ z_j\}$ ($\rho_2 = \{(z_i, z_j) : z_j \succ z_i\}$).

Теорема 3.1. Пусть заданы множества (наборы, тесты) признаков $\Theta_1 = \{z_1, z_2, \dots, z_{g-1}\}$ и $\Theta_2 = \{z_1, z_2, \dots, z_{g-1}, z_g\}$ ($\Theta_1 \subset \Theta_2$). Бинарные отношения (предпочтения) $z_i \rho_1 z_j, z_i \rho_2 z_j, z_i, z_j \in \Theta_1 \subset \Theta_2, i \neq j$, индуцированные на множествах Θ_1 и Θ_2 посредством применения классической процедуры МАИ, в общем случае не совпадают.

Теорема 3.2. Пусть для МПС альтернатив $\mathbf{A}' = \|a'_{ij}\|_{(g-1) \times (g-1)}$ ($a'_{ij} = w'_i/w'_j$) и $\mathbf{A} = \|a_{ij}\|_{g \times g}$ ($a_{ij} = w_i/w_j$) множеств Θ_1 и Θ_2 , соответственно, выполнены свойства (3.1) МПС a)-d). Тогда при выполнении условий:

$$1) a'_{ij} = a_{ij}, \text{ для } i, j < g; 2) (a'_{ij})^g > 1 \left((a'_{ij})^g < 1 \right)$$

бинарные отношения (предпочтения) $z_i \succ z_j$ ($z_i \prec z_j$), $z_i, z_j \in \Theta_1 \subset \Theta_2, i \neq j$, индуцированные на множествах Θ_1 и Θ_2 посредством применения стандартной процедуры метода анализа иерархий, совпадают.

Теорема 3.3. Пусть заданы множества (наборы, тесты) альтернатив Θ_1 и Θ_2 , $\Theta_1 \subset \Theta_2$. Бинарные отношения (предпочтения) $z_i \rho_1 z_j, z_i \rho_2 z_j, z_i, z_j \in \Theta_1 \subset \Theta_2, i \neq j$, индуцированные посредством применения классической процедуры МАИ на множестве Θ_1 и модифицированной процедуры ММАИ на множестве Θ_2 , совпадают.

Из табл.3.1 следует, что функция, на основе которой принимается решение в ММАИ (назовем ее функцией относительного сходства i -й альтернативы по всей совокупности альтернатив и критериев), имеет вид:

$$f_i = \frac{1}{g} \sum_{j=1}^g w_{ij}(i) \left(f_i = \left[\prod_{j=1}^g w_{ij}(i) \right]^{1/g} \right),$$

$$w_{ij}(i) = \sum_{s=1}^k c_s w_{ij}^s(i) = \sum_{s=1}^k c_s \frac{w_i^s}{w_i^s + w_j^s}, \quad w_i^s = \left(\prod_{l=1}^g a_{il}^s \right)^{1/g} \left[\sum_{k=1}^g \left(\prod_{l=1}^g a_{kl}^s \right)^{1/g} \right]^{-1},$$

величины $a_{kl}^s = w_k^s / w_l^s$ - элементы МПС по s -му критерию МАИ.

Сформулируем достаточное условие парето-оптимальности функции относительного сходства в терминах множества выбираемых решений-альтернатив X и множества, оптимального по Парето $P_f(X)$.

Теорема 3.4. Если функция относительного сходства в точке x' достигает максимального значения на множестве выбираемых решений, то x' входит в парето-оптимальное множество.

Доказательства теорем 3.1-3.4 - в приложении 3.

Следует также отметить, что свойства (3.1) МПС a - d) являются условиями применения МАИ и считаются по умолчанию выполненными. Невыполнение указанных свойств приводит к некорректному использованию МАИ (см. утверждение 3.1).

3.4. Иллюстративный пример применения ММАИ

Изложенный выше метод ММАИ, несмотря на простоту, нуждается в численной иллюстрации. Покажем этапы ММАИ на данных примера 3.2.

Этап 1. Пусть МПС альтернатив (признаков) составлена по каждому из двух данных равновесных критериев и посчитаны весовые коэффициенты. Матрицы парных сравнений \mathbf{B}_I^1 и \mathbf{B}_{II}^1 представлены для удобства в расширенном виде, включая и оценки собственных векторов w_i^j и их нормализованные значения $w_{iнорм}^j$ (верхний индекс указывает номер критерия, нижний – номер этапа). Вычисления проведем с точностью до 3-х знаков.

$$\mathbf{B}_I^1 = \begin{matrix} & z_1 & z_2 & z_3 & w_1^1 & w_{норм}^1 \\ \begin{bmatrix} 1.000 & 0.500 & 3.000 & 1.145 & 0.300 \\ 2.000 & 1.000 & 6.000 & 2.289 & 0.600 \\ 0.333 & 0.167 & 1.000 & 0.382 & 0.100 \end{bmatrix}; \end{matrix}$$

$$\mathbf{B}_I^2 = \begin{bmatrix} z_1 & z_2 & z_3 & w_1^2 & w_{1норм}^2 \\ 1.000 & 4.000 & 0.500 & 1.260 & 0.308 \\ 0.250 & 1.000 & 0.125 & 0.315 & 0.077 \\ 2.000 & 8.000 & 1.000 & 2.520 & 0.615 \end{bmatrix}.$$

По 2-м равновесным критериям оценки альтернатив, следуя традиционному МАИ, будут равны: $w^0=(0.316, 0.352, 0.332)$. Согласно модифицированному МАИ далее следуют действия.

Этап 2. По двум критериям формируем МПС $\mathbf{B}_{II}^1 = \|w_{ij}^1\| = \|(w_i^1, w_j^1)\|$ и $\mathbf{B}_{II}^2 = \|w_{ij}^2\| = \|(w_i^2, w_j^2)\|$:

$$\mathbf{B}_{II}^1 = \begin{bmatrix} (1.000;1.000) & (0.300;0.600) & (0.300;0.100) \\ (0.600;0.300) & (1.000;1.000) & (0.600;0.100) \\ (0.100;0.300) & (0.100;0.600) & (1.000;1.000) \end{bmatrix};$$

$$\mathbf{B}_{II}^2 = \begin{bmatrix} (1.000;1.000) & (0.308;0.077) & (0.308;0.615) \\ (0.077;0.308) & (1.000;1.000) & (0.077;0.615) \\ (0.615;0.308) & (0.615;0.077) & (1.000;1.000) \end{bmatrix}.$$

Далее нормализуем по каждому критерию каждую пару элементов сравнения альтернатив и получаем относительные локальные ВКА (по формулам 2-й строки, 3-го и 4-го столбцов в табл.1), например:

$$w_{12}^1 = \left(\frac{0.300}{0.300 + 0.600}; \frac{0.600}{0.300 + 0.600} \right) = (0.333; 0.667), \text{ и оформляем их в матрицы нормализованных локальных ВКА по каждому критерию:}$$

$$\mathbf{B}_{II}^1 = \begin{bmatrix} (0.500;0.500) & (0.333;0.667) & (0.750;0.250) \\ (0.667;0.333) & (0.500;0.500) & (0.857;0.143) \\ (0.250;0.750) & (0.143;0.857) & (0.500;0.500) \end{bmatrix};$$

$$\mathbf{B}_{II}^2 = \begin{bmatrix} (0.500;0.500) & (0.800;0.200) & (0.333;0.667) \\ (0.200;0.800) & (0.500;0.500) & (0.111;0.889) \\ (0.667;0.333) & (0.889;0.111) & (0.500;0.500) \end{bmatrix}.$$

Этап 3. Формируем обобщенную матрицу $\mathbf{W} = \|w_{ij}\| = \|(w_{ij}(i), w_{ij}(j))\|$, где векторы $(w_{ij}(i), w_{ij}(j))$ – локальные ВКА z_i, z_j уровня 2 относительно *всей совокупности критериев* сравнения, компоненты которых находим по формулам 4-й строки табл.1, например,

$$(u_{12}^1; u_{12}^2) = \left(\frac{1}{2}(0.333 + 0.800); \frac{1}{2}(0.667 + 0.200) \right) = (0.567; 0.433).$$

$$\mathbf{W} = \begin{bmatrix} (0.500; 0.500) & (0.567; 0.433) & (0.542; 0.458) \\ (0.433; 0.567) & (0.500; 0.500) & (0.484; 0.516) \\ (0.458; 0.542) & (0.516; 0.484) & (0.500; 0.500) \end{bmatrix}.$$

Этап 4. Глобальные значения ВКА считаем по формуле: $V_i^{(1)} = \sum_{j=1}^g w_{ij}$, и

затем нормализуем: $V = (0.357; 0.315; 0.328)$.

Таким образом, согласно модифицированной процедуре ВКА первых двух альтернатив будут равны: 0.357; 0.315, что согласуется с оценками, полученными по 2-м критериям в первоначальном наборе альтернатив.

Продолжая рассмотрение примера 3.1, получим следующие относительных весовых коэффициентов:

$$\mathbf{B}'_1 = \begin{bmatrix} (7; 7) & (7; 10) & (7; 5) \\ (10; 7) & (10; 10) & (10; 5) \\ (5; 7) & (5; 10) & (5; 5) \end{bmatrix}; \quad \mathbf{B}'_2 = \begin{bmatrix} (15; 15) & (15; 10) & (15; 20) \\ (10; 15) & (10; 10) & (10; 20) \\ (20; 15) & (20; 10) & (20; 20) \end{bmatrix};$$

$$\mathbf{B}_1 = \begin{bmatrix} (0.500; 0.500) & (0.412; 0.588) & (0.583; 0.417) \\ (0.588; 0.412) & (0.500; 0.500) & (0.667; 0.333) \\ (0.417; 0.583) & (0.333; 0.667) & (0.500; 0.500) \end{bmatrix};$$

$$\mathbf{B}_2 = \begin{bmatrix} (0.500; 0.500) & (0.600; 0.400) & (0.429; 0.571) \\ (0.400; 0.600) & (0.500; 0.500) & (0.333; 0.667) \\ (0.571; 0.429) & (0.667; 0.333) & (0.500; 0.500) \end{bmatrix}.$$

Ненормализованные оценки собственных векторов, согласно ММАИ, будут равны: $w^1 = (0.493, 0.580, 0.411)$, $w^2 = (0.505, 0.405, 0.575)$, соответственно. Итоговый весовой вектор по равновесным критериям равен: $w = (0.336, 0.332, 0.332)$, что отвечает верному выбору.

Приведенные примеры свидетельствуют о том, что, по-видимому, не существует некоей универсальной функции (полезности) скаляризации критериев с целью выбора в качестве наилучшего произвольного элемента из оптимального множества Парето. Обсуждаемая модификация известного метода позволяет надеяться, что на основе вышеопределенной функции относительного сходства возможно непротиворечивое упорядочение анализируемых элементов (альтернатив, признаков). На основе данной модификации метода парных сравнений, успешно решен ряд прикладных задач (например, [91]).

Выводы. В разделе рассмотрены свойства модификации метода анализа иерархий, расширяющие возможности классического метода, а именно: а) доказа-

но, что сохранение отношения предпочтения при изменении множества сравниваемых альтернатив МАИ не гарантирует (теорема 3.1); б) доказано, что модификация метода анализа иерархий гарантирует сохранение отношения предпочтения при изменении множества сравниваемых альтернатив МАИ (теорема 3.2); в) доказана парето-оптимальность выбираемых решений по модифицированному методу анализа иерархий (теорема 3.4).

Метод может быть применен для решения слабоструктурированных задач, в интеллектуальных системах поддержки принятия решений, при решении задачи прогнозирования. Как известно, одним из этапов в последовательности шагов прогнозирования свойств (поведения) объекта является оценка нововведения (версии, альтернативы) по критериям с последующим групповым упорядочением нововведений. Нелинейная модификация ММАИ свободна от некорректностей в решении многокритериальных слабоформализованных задач, которые могут привести к ошибочному решению.

Метод внедрен в интеллектуальную инструментальную систему ИМСЛОГ [24] и прошел успешную апробацию.

На основе метода возможен выбор наиболее информативной системы признаков, метода (алгоритма) распознавания (альтернативы), а также формирование коллективного решения относительно ранга распознаваемого состояния ДО решением многокритериальной задачи на основе имеющихся данных по частным критериям: качество распознавания состояний, время распознавания, полученных на контрольной выборке. Качество распознавания количественно выражается в числе правильно распознанных объектов (значений ВР) относительно общего числа предъявленных к распознаванию. Отличительной особенностью применяемого для этой цели корректной модификации известного метода парных сравнений является возможность корректного анализа динамически изменяющихся сравниваемых наборов альтернатив и гарантированное нахождение оптимума (наилучшее решение будет доставлять максимум используемой свертки критериев – функции относительного сходства).

Результаты, представленные в 3-й главе, опубликованы в работах [11, 16, 72, 76, 79, 91, 96, 174].

«...любое детальное исследование неизбежно требует перехода к дискретному описанию. Справедливость этого тезиса становится особенно очевидной, когда мы начинаем анализировать процессы с использованием ЭВМ. Первый шаг такого анализа – это всегда переход к дискретному представлению изучаемых моделей... Успехи именно в этой области определяют качественный прорыв в будущее» /Академик Н.Н. Моисеев/

Глава 4. МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ В ДИСКРЕТНОМ ПОДХОДЕ К РАСПОЗНАВАНИЮ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА

Прежде чем изложить общую схему тестового распознавания [55, 108, 175] применительно к решению задачи идентификации состояний сложного ДО, дадим постановку задачи распознавания состояний ДО в удобной для исследования форме (см. гл. 1).

Постановка задачи. Пусть имеется некоторая динамическая система, которая может находиться в одном из I состояний, информация о которых содержится в амплитудах сигналов, измеряемых g датчиками в дискретные моменты времени t и являющихся случайными значениями временного ряда. Задача состоит в том, чтобы, имея обучающую выборку (экспериментальные данные показаний периодически опрашиваемых датчиков при разных состояниях ДО) и анализируя экспериментальные показания датчиков в реальном времени, определить метку (номер) состояния ДО.

4.1. Основные понятия и определения

В настоящее время техника анализа признаков описаний используется в методах оптимизации многопараметрических моделей распознавания. Выявление информативных фрагментов описаний обучающих объектов дает возможность проводить качественный анализ исходной информации, например, с целью классификации признаков и уменьшения их числа.

Известно [52], что центральной проблемой дискретного подхода к задачам распознавания является поиск информативных фрагментов признаков описаний объектов. При поиске информативных фрагментов используется аппарат логических функций, в частности методы преобразования нормальных форм булевых функций, а также теория покрытий булевых и целочисленных матриц. Основные модели дискретных (логических) процедур распознавания и вопросы, связанные со сложностью их реализации, рассмотрены в ряде работ (например, [22, 23, 52, 55, 88, 175]). Формализуем используемые далее понятия, опираясь на работы [52, 55, 156].

Стандартная постановка задачи распознавания заключается в следующем. Исследуется некоторое множество объектов O . Объекты этого множества описываются системой признаков $\{z_1, \dots, z_n\}$. Известно, что множество O представимо в виде объединения непересекающихся подмножеств (классов) $\{\Omega_1, \dots, \Omega_r\}$. Имеется конечный набор объектов $\{Y_1, \dots, Y_m\}$ из O , о которых известно, каким классам они принадлежат (это прецеденты или обучающие объекты).

Требуется по предъявленному набору значений признаков, т.е. описанию некоторого объекта Y из O (не входящего в обучающее множество), о котором, вообще говоря, неизвестно, какому классу он принадлежит, определить этот класс.

Под закономерностью здесь понимается подобласть признакового пространства, содержащая объекты одного класса (полная закономерность) или преимущественно одного класса (частичная закономерность).

При построении классических дискретных процедур вводится понятие элементарного классификатора [52], под которым понимается фрагмент описания обучающего объекта.

Определение 4.1. Пусть H – некоторый набор из r различных признаков вида $H = \{z_{j_1}, \dots, z_{j_r}\}$, и пусть Y – некоторый объект из O , $Y = (a_1, \dots, a_n)$, здесь a_j – значение признака z_j , $j = 1, 2, \dots, n$. Набор H выделяет в описании объекта Y фрагмент $\{a_{j_1}, \dots, a_{j_r}\}$. Если объект Y принадлежит обучающему множеству, то набор $\{a_{j_1}, \dots, a_{j_r}\}$ называют элементарным классификатором (ЭК), порожденным признаками из H и обозначают (Y, H) .

Близость объектов $Y' = (a'_1, a'_2, \dots, a'_g)$ и $Y'' = (a''_1, a''_2, \dots, a''_g)$ из класса $\cup \subset (\Omega_1, \dots, \Omega_r)$ по набору признаков $H = \{z_{j_1}, \dots, z_{j_r}\}$ будем оценивать величиной:

$$B(Y', Y'', H) = \begin{cases} 1, & a'_{j_t} = a''_{j_t}, t \in \{1, 2, \dots, r\} \\ 0, & \text{иначе} \end{cases}, \quad (4.1)$$

если признаки целочисленные, и величиной:

$$B(Y', Y'', H) = \begin{cases} 1, & |a'_{j_t} - a''_{j_t}| < e_{j_t}, t \in \{1, 2, \dots, r\} \\ 0, & \text{иначе} \end{cases}, \quad (4.2)$$

если признаки вещественнозначные (e_m – точность измерения признака z_m).

Таким образом, объекты Y' и Y'' близки по набору признаков H ($B(Y', Y'', H) = 1$) тогда и только тогда, когда имеет место совпадение фрагментов (Y', H) и (Y'', H) .

Определение 4.2. Фрагмент описания объекта Y' из класса ν вида (Y', H) назовем представительным набором для ν , если для любого обучающего объекта Y'' , не принадлежащего классу ν , имеет место $B(Y', Y'', H) = 0$.

Определение 4.3. Фрагмент описания объекта Y' из класса ν вида (Y', H) назовем тупиковым представительным набором для ν , если выполнены два условия: 1) для любого обучающего объекта $Y'' \notin \nu$ имеет место $B(Y', Y'', H) = 0$; 2) для любого набора $H' \in H$, найдется обучающий объект $Y'' \notin \nu$, для которого $B(Y', Y'', H) = 1$. Тупиковый представительный набор является неизбыточным (при сжатию он теряет способность отличать порождающий его объект от некоторых объектов из других классов).

Определение 4.4. Набор признаков H назовем тестом, если для любых обучающих объектов Y' и Y'' , принадлежащих разным классам, выполнено $B(Y', Y'', H) = 0$. Другими словами, тест - это набор признаков, по которому различаются любые два объекта из разных классов.

Понятно, что если набор признаков H_1 - тест, то любой набор признаков H_2 такой, что $H_1 \subseteq H_2$, также является тестом. Однако, если объекты близки по H_2 , то они будут близки и по H_1 , если же два объекта близки по набору столбцов H_1 , то они не всегда будут близки по H_2 . В этом смысле более короткие тесты обладают большей информативностью, и есть смысл строить именно тупиковые тесты.

Общая схема принятия решения в распознающем алгоритме следующая²². На этапе обучения для каждого класса $\nu \subset (\Omega_1, \dots, \Omega_l)$ строится некоторое множество элементарных классификаторов V_ν с заданными свойствами. На этапе распознавания осуществляется процедура голосования.

Распознаваемый объект Y сравнивается с каждым обучающим объектом Y_i по каждому тесту. В простейшем варианте считается, что объект Y получает голос за принадлежность классу ν , если $Y_i \in \nu$ и описания объектов Y и Y_i совпадают по множеству признаков H (в этом случае $B(Y, Y_i, H) = 1$). Для каждого класса $\nu \subset (\Omega_1, \dots, \Omega_l)$ вычисляется оценка принадлежности $G(Y, \nu)$ объекта Y классу ν , которая имеет вид:

$$G(Y, \nu) = \frac{1}{|V_\nu|} \sum_{(Y_i, H) \in V_\nu} B(Y, Y_i, H), \quad (4.3)$$

где $|X|$ - мощность множества X .

²² Впервые алгоритм голосования по представительным наборам описан в работе: Баскакова Л.В., Журавлёв Ю.И. Модель распознающих алгоритмов с представительными наборами и системами опорных множеств // Ж. вычисл. матем. и матем. физ. - 1981. - Т. 21. - №5. - С. 1264-1275.

Если $G(Y, \nu_s) = \max_{1 \leq j \leq I} G(Y, \nu_j)$, $G(Y, \nu_s) \neq G(Y, \nu_u)$, $s \neq u$, то объект Y относят к классу ν_s ; если $\max_{1 \leq j \leq I} G(Y, \nu_j) = G(Y, \nu_s) = G(Y, \nu_u)$, $s \neq u$, то происходит отказ от распознавания.

Оценивание качества алгоритма распознавания производится, как правило, на основе процедуры скользящего контроля.

Матричное представление данных и знаний. В интеллектуальных системах выявления закономерностей и принятия решений, основанных на тестовых методах распознавания образов, используется модель представления знаний в виде матрицы «объект-признак».

Нижеприведенный метод и реализующий его алгоритм распознавания состояний ДО основан на матричной модели представления данных и знаний [23, 175], включающей матрицу описаний (**Q**) объектов в пространстве характеристических признаков и матрицу различий (**R**) объектов в пространстве классификационных признаков (А.Е. Янковская). Преимущество данной модели определяется возможностью использования объектов и методов дискретной математики, параллельной обработкой информации и применением одних и тех же ключевых алгоритмов на разных этапах реализации интеллектуальной системы тестового распознавания, начиная от анализа знаний на непротиворечивость и кончая принятием решений. На рис. 4.1 приведены примеры матриц описаний **Q**, различий **R** и тестов **T** [23, 175].

$$\begin{array}{c}
 \begin{array}{cccccccccccc}
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
 \begin{array}{l}
 1 \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \\
 2 \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\
 3 \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & - & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\
 4 \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 5 \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \\
 6 \begin{bmatrix} 1 & - & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{bmatrix} \\
 7 \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & - & 0 \end{bmatrix} \\
 8 \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}
 \end{array}
 \end{array}
 \end{array}
 ; \mathbf{R} = \begin{array}{c}
 \begin{array}{c}
 \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}
 \end{array}
 \end{array}
 ; \mathbf{T} = \begin{array}{c}
 \begin{array}{cccccccccccc}
 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\
 1 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 2 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \\
 3 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 4 \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 5 \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 6 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}
 \end{array}
 \end{array}
 \end{array}$$

Рис. 4.1. Матрицы описаний **Q**, различий **R** и тестов **T**.

Элемент q_{ij} матрицы **Q** задает значение j -го признака для i -го объекта. Строка q_i матрицы **Q** сопоставляется объекту Y_i ($i \in \{1, \dots, N\}$), где N – число обучающих объектов. В случае, если у какого-либо объекта в матрице **Q** значение признака отмечено символом «-», то считается, что значение соответст-

вующего характеристического признака безразлично (признак может принимать как нулевые, так и единичные значения, а в случае k -значных признаков – любые целочисленные значения из заданного интервала значений признака).

Строки матрицы различений \mathbf{R} сопоставляются одноименным строкам матрицы \mathbf{Q} , столбцы – классификационным признакам, определяющим различные механизмы разбиения объектов на классы эквивалентности (механизмы классификации). Элемент r_{ij} матрицы \mathbf{R} задает принадлежность i -го объекта одному из выделенных классов по j -му механизму классификации. Для указания факта принадлежности объекта классу используется номер этого класса при соответствующем механизме классификации.

Множество всех неповторяющихся строк матрицы \mathbf{R} сопоставлено множеству выделенных образов. Элементами образа являются объекты, представленные строками матрицы \mathbf{Q} , сопоставленными одинаковым строкам матрицы \mathbf{R} .

Если имеется единственный механизм классификации, матрица различений вырождается в столбец, что соответствует традиционному представлению знаний в задачах распознавания образов [55].

Реакция матрицы описаний на тест есть совокупность строк матрицы описаний, в которых значение признаков, входящих в тест, совпадает со значениями аналогичных признаков исследуемого объекта.

Задача распознавания состоит в определении по матрицам \mathbf{Q} и \mathbf{R} образа, которому принадлежит заданный совокупностью признаков исследуемый объект, как правило, не входящий в обучающую выборку.

Для решения задачи распознавания вводится понятие закономерности в данных и знаниях, используемое в создании ряда интеллектуальных систем, например, [39, 40] (гл. 1, определение 1.16).

Будем использовать понятие зависимости признаков, отличное от приведенного в [175].

Назовем признаки зависимыми, если имеется хотя бы одна пара объектов из разных образов, различаемая этими признаками.

Совокупность признаков, различающих все пары объектов из разных образов (классов, при каждом механизме классификации), назовем диагностическим тестом (далее слово «диагностический» будем опускать). Другими словами, набор столбцов таблицы \mathbf{Q} назовем тестом, если в подтаблице, образованной этим набором, любые две строки, описывающие объекты из разных классов, различны.

Как было отмечено выше, тест называют тупиковым [55] (безызбыточным [175]), если исключение любого признака из теста нарушает его свойство быть тестом.

Совокупность тупиковых тестов представим матрицей \mathbf{T} , строки которой сопоставлены тестам, а столбцы – характеристическим признакам. Строка T_i матрицы \mathbf{T} (рис. 4.1) соответствует i -му тупиковому (безызбыточному) тесту, а столбцы –

признакам из совокупности $Z = (z_1, z_2, \dots, z_M)$, каждый из которых содержится хотя бы в одном тесте, $T_{ij} = 1 \leftrightarrow z_j \in \{z_{i_1}, z_{i_2}, \dots, z_{i_h}, \dots, z_{i_g}\}$ ($i_h \in \{1, \dots, M\}$, g – количество признаков в тесте), M – количество признаков во всех тупиковых тестах, представленных строками матрицы \mathbf{T} .

Множество обязательных признаков (входящих во все тупиковые тесты [23, 175]) называют ядром всех диагностических тестов, поскольку исключение любого признака из ядра нарушает свойство каждого из тестов быть тестом.

Заметим, что количество столбцов в матрицах \mathbf{Q} и \mathbf{T} , как правило, не совпадает, поскольку в матрицу \mathbf{T} включены только столбцы, сопоставленные характеристическим признакам, вошедшим во все тесты, соотнесенными со строками матрицы \mathbf{T} . Однако для простоты дальнейшего изложения примем число столбцов в обеих матрицах равным M .

При $q_{ij} \in \{0, 1, \langle - \rangle\}$ вводятся следующие определения.

Два объекта считаются различимыми, если хотя бы один характеристический признак в описании одного объекта принимают значение 1 (0), а в описании другого – инверсное, то есть значение 0 (1).

Под весовым коэффициентом признака понимается числовая оценка его различающей способности [175] или информационный вес признака [55].

Формализм мультимножеств используется не только при изложении методов определения весовых коэффициентов признаков и тестов, но и в гл. 5 при изложении постановки задачи выбора оптимальной системы градаций квантования как системы признаков состояний ДО (определения 5.1, 5.2).

4.2. Формализм мультимножеств

Мультимножеством A , порожденным обычным множеством $U = \{x_1, x_2, \dots\}$, все элементы которого различны, называется совокупность групп элементов вида $A = \{k_A(x) \cdot x \mid x \in U, k_A(x) \in \mathbb{Z}_+\}$. Здесь $k_A: U \rightarrow \mathbb{Z}_+ = \{0, 1, 2, \dots\}$ называется функцией числа экземпляров мультимножества, определяющей кратность вхождения элемента $x_i \in U$ в мультимножество A , что обозначено символом «•».

Если $k_A(x) = \chi_A(x)$, где $\chi_A(x) = 1$ при $x \in A$ и $\chi_A(x) = 0$ при $x \notin A$, то мультимножество A становится обычным множеством.

Если все мультимножества семейства $A = \{A_1, A_2, \dots\}$ образуются из элементов множества U , то U называется доменом для семейства A , а множество $\text{Supp}A = \{x \mid x \in U, \chi_{\text{Supp}A}(x) = \chi_A(x)\}$ – опорным множеством или носителем мультимножества A .

Мощность мультимножества $|A| = \sum_x k_A(x)$ определяется как общее число экземпляров всех его элементов; размерность мультимножества $/A/ = \sum_x \chi_A(x) = |\text{Supp}A|$ – как общее число различных элементов.

Максимальное значение функции кратности $\text{hgt}A = \max_{x \in G} k_A(x)$ называется высотой, а элемент $x_{A^*} = \arg \max_{x \in G} k_A(x)$ – пиком мультимножества A .

Мультимножество называется пустым \emptyset , если $k_{\emptyset}(x)=0$, и максимальным Z , если $k_Z(x) = \max_{A \in \mathcal{A}} k_A(x), \forall x \in U$.

Над мультимножествами выполняются традиционные теоретико-множественные операции, такие как объединение, пересечение, вычитание, дополнение, симметрическая разность, прямое произведение, и ряд новых операций: сложение, умножение, умножение на число, а также линейные комбинации операций. Приведем некоторые основные операции над мультимножествами:

- объединение $A \cup B = \{k_{A \cup B}(x) \bullet x \mid k_{A \cup B}(x) = \max(k_A(x), k_B(x))\}$;
- пересечение $A \cap B = \{k_{A \cap B}(x) \bullet x \mid k_{A \cap B}(x) = \min(k_A(x), k_B(x))\}$;
- арифметическое сложение $A + B = \{k_{A+B}(x) \bullet x \mid k_{A+B}(x) = k_A(x) + k_B(x)\}$;
- арифметическое вычитание $A - B = \{k_{A-B}(x) \bullet x \mid k_{A-B}(x) = k_A(x) - k_B(x)\}$;
- симметрическая разность $A \Delta B = \{k_{A \Delta B}(x) \bullet x \mid k_{A \Delta B}(x) = |k_A(x) - k_B(x)|\}$;
- дополнение $\bar{A} = Z - A = \{k_{\bar{A}}(x) \bullet x \mid k_{\bar{A}}(x) = k_Z(x) - k_A(x)\}$;
- умножение на число $h \bullet A = \{k_{h \bullet A}(x) \bullet x \mid k_{h \bullet A}(x) = h \cdot k_A(x), h \in \mathbb{Z}_+\}$;
- арифметическое умножение $A \bullet B = \{k_{A \bullet B}(x) \bullet x \mid k_{A \bullet B}(x) = k_A(x) \cdot k_B(x)\}$;
- арифметическая n -ая степень $A^n = \{k_{A^n}(x) \bullet x \mid k_{A^n}(x) = (k_A(x))^n\}$;
- прямое произведение $A \times B = \{k_{A \times B}(x_i, x_j) \mid k_{A \times B} = k_A(x_i) \cdot k_B(x_j), x_i \in A, x_j \in B\}$.

Более подробную информацию и примеры применения формализма мультимножеств можно найти в работах А.Б. Петровского (например, [131]), им же впервые введены следующие метрики (псевдометрики), на основе которых определяются различные классы метрических пространств мультимножеств (A, d) :

$$\begin{aligned} d_{1p}(A, B) &= [m(A \Delta B)]^{1/p}; \quad d_{2p}(A, B) = [m(A \Delta B)/m(Z)]^{1/p}; \\ d_{3p}(A, B) &= [m(A \Delta B)/m(A \cup B)]^{1/p}, \end{aligned} \quad (4.4)$$

где p – целое число, m – мера мультимножества, действительная неотрицательная функция, заданная на алгебре мультимножеств.

Перечислим некоторые способы задания меры мультимножеств: мощность мультимножества $m(A) = |A| = \sum_i k_A(x_i)$, линейная комбинация функций кратности $m(A) = \sum_i w_i k_A(x_i)$, $w_i > 0$. В последнем случае метрики (4.4) имеют вид:

$$d_{1p}(A, B) = \left(\sum_{x_i \in U} w_i |k_A(x_i) - k_B(x_i)| \right)^{1/p}; \quad (4.5)$$

$$d_{2p}(A, B) = \left(\sum_{x_i \in U} w'_i |k_A(x_i) - k_B(x_i)| \right)^{1/p}, \quad w'_i = w_i / \sum_{j=1}^h w_j k_Z(x_j); \quad (4.6)$$

$$d_{3p}(A, B) = \left(\frac{\sum_{x_i \in U} w_i |k_A(x_i) - k_B(x_i)|}{\sum_{x_i \in U} w_i \max[k_A(x_i), k_B(x_i)]} \right)^{1/p}, w_i > 0. \quad (4.7)$$

Прокомментируем содержательный смысл метрик (4.5)- (4.7).

Основная метрика $d_{1p}(A, B) = [m(A \Delta B)]^{1/p}$ (модуль симметрической разности) является метрикой типа Хемминга, традиционно используемой во многих приложениях.

Полностью усредненная метрика $d_{2p}(A, B) = [m(A \Delta B)/m(Z)]^{1/p}$ характеризует различие между двумя мультимножествами A и B , отнесенное к расстоянию, максимально возможному в исходном пространстве.

Локально усредненная метрика $d_{3p}(A, B) = [m(A \Delta B)/m(A \cup B)]^{1/p}$ задает различие, отнесенное к максимально возможной «общей части» $A \cup B$ только этих двух мультимножеств в исходном пространстве. Функции $d_{2p}(A, B)$ и $d_{3p}(A, B)$ удовлетворяют условию нормировки $0 \leq d(A, B) \leq 1$. Функция $d_{3p}(A, B)$ не определена для $A = B = \emptyset$, поэтому по определению принимается $d_{3p}(\emptyset, \emptyset) = 0$.

Ниже даны новые меры относительной важности мультимножеств, позволившие сравнивать взаимозависимые признаки и определять их весовые коэффициенты.

4.3. Алгоритм тестового распознавания состояния ДО

Отметим, что применение дискретного подхода оказывается во многих случаях сложным в силу чисто вычислительных трудностей переборного характера, возникающих на этапе поиска информативных фрагментов описаний объектов [52]. При числе признаков, равным M , число непустых подмножеств составляет 2^{M-1} , и прямой перебор всех подмножеств оказывается невозможным уже при M порядка 20 даже при современных вычислительных возможностях. В силу экспоненциального роста числа фрагментов при возрастании размерности описаний, решение проблемы только за счет повышения производительности вычислительной техники нереально, поэтому применение дискретных методов весьма затруднительно при большом объеме обучающей информации [50, 52].

Непосредственное применение тестового распознавания для решения поставленной в начале главы 4 задачи невозможно [50] при большом числе признаков (в силу NP -полноты задачи). Однако вышеизложенный алгоритм обобщенных эталонов (алгоритм 2.1) позволяет не только существенно ее сократить (в десятки и более раз), но и повысить репрезентативность выборки, представленной в виде «объект-признак».

Выходом каждого m -го датчика z_m ($m = \overline{1, g}$) являются дискретные значения индикаторных показателей, косвенно характеризующих состояния ДО.

Сформируем из исходных данных, полученных наблюдением за поведением ДО в разных состояниях, обучающую матрицу типа объект-признак \mathbf{Q} объектов в пространстве характеристических признаков (показаний датчиков) и матрицу различий \mathbf{R} объектов в пространстве классификационных признаков.

По матрицам \mathbf{Q} и \mathbf{R} каким-либо из методов [22, 23, 52, 55, 88, 175] построим все (часть) тупиковых тестов (рис. 4.1), представленные матрицей тестов \mathbf{T} , строки которой сопоставлены тестам, а столбцы – характеристическим признакам, и определим число различающих пар «образ-образ» по каждому признаку, где под образом будем понимать состояние ДО. Для корректного принятия решения в тестовом распознавании используется понятие весовых коэффициентов признаков (ВКП) и весовых коэффициентов тестов (ВКТ), под которыми понимается числовая оценка их различающей способности. Приведение исходной вышеставленной прикладной задачи к стандартной задаче распознавания образов заключается в следующем.

Исследуется некоторое множество объектов Ω . Объекты этого множества описываются системой признаков (показаний датчиков) $Z = \{z_m, m = \overline{1, g}\}$. Известно, что множество Ω представимо в виде объединения непересекающихся подмножеств-классов (состояний) $\Omega = (\Omega_1, \dots, \Omega_l)$. Имеется конечный набор объектов (Y_1, \dots, Y_n) из Ω , о которых известно, каким классам они принадлежат (это прецеденты или обучающие объекты - реализации СП, а также обобщенные эталоны (G -эталон)). Предполагается, что описания объектов из разных классов (состояний) различаются.

Требуется по предъявленному набору значений признаков $Y = (a_1, \dots, a_g)$ (a_m – значение признака $z_m, m = \overline{1, g}$), т.е. описанию некоторого объекта Y из Ω (не входящего в обучающее множество), определить этот класс (состояние ДО). В случае $Y \in \{Y_1, \dots, Y_n\}$ (обучающее множество) фрагмент $\{a_{j_1}, \dots, a_{j_g}\}$ – элементарный классификатор, обозначим его через (Y, Z) .

Пусть величины U_Z – множество всех ЭК, $U_Z = \{(Y, Z) | Y \in \{Y_1, \dots, Y_m\}\}$; U_ν ($\nu \subset (\Omega_1, \dots, \Omega_l)$) – множество всех ЭК из U_Z , порождаемых обучающими объектами из класса ν , $U_\nu = \{(Y, Z) | Y \in \{Y_1, \dots, Y_m\}, Y \in \nu\}$, соответственно. Поскольку более информативными считаются тупиковые тесты [55, 56], т.е. минимальные наборы признаков, позволяющие безошибочно разделять обучающий материал на классы, то далее в качестве ЭК будем рассматривать короткие тупиковые наборы (тесты), алгоритмы для выбора которых [22, 23], изложены во многих работах²³ (см. в приложении П4.1 пример эвристического алгоритма ус-

²³ Дюкова Е.В. Дискретные (логические) процедуры распознавания: принципы конструирования, сложность реализации и основные модели. Учебное пособие для студентов математических факультетов педвузов. - М.: Прометей, 2003. – 29 с.

коренного поиска всех кратчайших столбцовых покрытий (тупиковых тестов) двоичной матрицы на основе метода ветвей и границ, использующего оригинальную разметку столбцов, позволяющую отсекать неоптимальные решения).

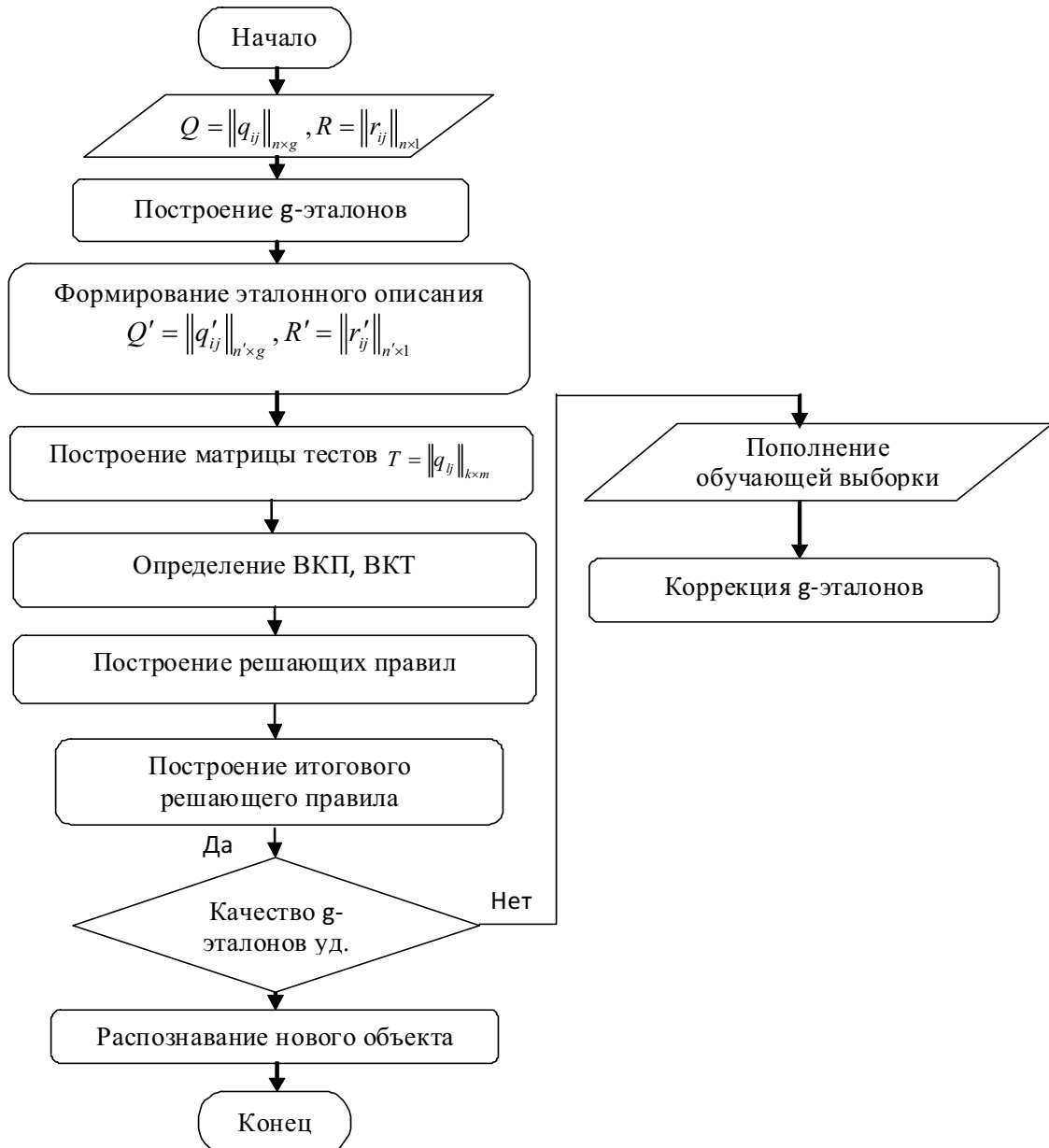


Рис. 4.2. Блок-схема тестового распознавания состояния ДО с предобработкой исходных данных с целью получения информативных взвешенных g-эталонов

Ниже излагается алгоритм 4.1 тестового распознавания состояния ДО соответственно блок-схеме (рис. 4.2).

Fulkerson D.R., Nemhauser G.L., Trotter L.E. Two computationally difficult set covering problems that arise in computing the 1-width of incidence matrices of steiner triple systems// Mathematical Programming Study 2 (1974) 72-81. North-Holland Publishing Company.

Без ограничения общности алгоритм 4.1 излагается для 2-х состояний (образов, классов).

Алгоритм 4.1. Модификация тестового распознавания состояния ДО

Вход: Обучающая выборка $Q = \|q_{jk}\|_{n \times m}$, $R = \|r_{jl}\|_{n \times 1}$, или совокупность объектов $U = \{Y_1, Y_2, \dots, Y_n\}$, представленная значениями признаков $Y_j = (q_{j1}, \dots, q_{jm})$, $j = \overline{1, n}$, в виде размеченного временного ряда (без ограничения общности предположим, что объект (значение временного ряда) Y_j принадлежит одному из состояний Ω_1 или Ω_2 , при этом последний признак q_{jm} интерпретируется как момент времени t_j : $Y_j = Y(t_j)$, δ – параметр алгоритма, пороговое значение для сравнения коэффициентов относительного сходства g -эталонов и объектов для пополнения обучающей выборки.

Выход: Эталонное описание $Q' = \|q'_{jk}\|_{n' \times m}$, $R' = \|r'_{jl}\|_{n' \times 1}$; множество решающих правил (тестов T); итоговое решающее правило; номер класса распознаваемого объекта i_0 .

1: Формировать по алгоритму 2.1 обобщенные g -эталоны $U_G = \{G_1, G_2, \dots, G_n\}$, $n' \leq n$, $G_j = (q'_{j1}, \dots, q'_{jm}, w_j)$, где q'_{jk} – усредненное значение k -го признака z_k для j -го g -эталона G_j ($k = \overline{1, m}$), w_j – число объектов, вошедших в g -эталон G_j (весовой коэффициент g -эталона G_j).

2: Формировать эталонное описание $Q' = \|q'_{jk}\|_{n' \times m}$, $R' = \|r'_{jl}\|_{n' \times 1}$.

3: Построить множества решающих правил (число способов распознавания равно числу используемых для распознавания тестов, представленных матрицей тестов T) по одному из алгоритмов, например алгоритму П4.1 (приложение 4).

4: Проверить дискриминирующее качество эталонного описания на контрольной выборке. Если качество удовлетворительное, то переходим к шагу 6, в противном случае пополняем выборку по алгоритму шага 5, особенность которого – в учете весовых коэффициентов сформированных на шаге 1 обобщенных g -эталонов.

5: Осуществить пополнение обучающей выборки с коррекцией существующих эталонов и добавлением новых по следующей схеме.

5.1. Цикл по $l = \overline{1, l^*}$. Обозначим пополняемый обучающую выборку объект $Y_l^* = (q_{l1}^*, \dots, q_{lm}^*)$, где l^* – число дополнительных объектов.

5.2. Вычислить коэффициенты относительного сходства между пополняемыми обучающую выборку объектами $Y_l^* = (q_{l1}^*, \dots, q_{lm}^*)$, $l = \overline{1, l^*}$ и всеми g -

эталоны G'_j состояния Ω_1 по формуле: $\lambda_{lj} = w_j \left[\sum_{e=1}^g (q_{le}^* - q'_{je})^2 \right]^{-1}$. Отнесение l^* -го объекта к одному из существующих эталонов осуществить по следующему правилу 5.3-5.5.

5.3. Найти ближайший (к анализируемому объекту Y_l^*) g -эталон «своего» класса с номером $\alpha = \arg \max_{i=1, n', G'_i \subset \Omega_1} \lambda_{li}$, и ближайший g -эталон класса-конкурента с номером $\beta = \arg \max_{i=1, n', G'_i \not\subset \Omega_1} \lambda_{li}$.

5.4. Принять решение $Y_l^* \in G_\alpha$, и значения признаков g -эталона G_α пересчитать: $q'_{\alpha j} := \frac{q'_{\alpha j} \cdot w_\alpha + q_{lj}^*}{w_\alpha + 1}$, $w_\alpha := w_\alpha + 1$, если $\lambda_{l\alpha} - \lambda_{l\beta} \geq \delta$.

5.5. Принять решение $Y_l^* \in G_\beta$, и значения признаков g -эталона G_β пересчитать: $q'_{\beta j} := \frac{q'_{\beta j} \cdot w_\beta + q_{lj}^*}{w_\beta + 1}$, $w_\beta := w_\beta + 1$, если $\lambda_{l\alpha} - \lambda_{l\beta} < -\delta$.

5.6. Принять решение о появлении нового g -эталона $n' := n' + 1$, $G_{n'} := Y_l^*$, $w_{n'} := w_l$, если $(|\lambda_{l\alpha} - \lambda_{l\beta}| < \delta)$.

5.7. Конец цикла по $l = 1, l^*$. Положить $U_G = \{G_1, G_2, \dots, G_{n'}\}$

5.8. Вернуть $U_G = \{G_1, G_2, \dots, G_{n'}\}$.

5.9. Принять решение об удалении из U'_G объектов, веса которых меньше заданного порогового G'_u : $U'_G := U'_G \setminus G'_u$, $w'_i < w_0$ (см. рис. 2.4) с целью фильтрации выборки от «шума».

5.10. Выполнять шаг 2.

6: Для вычисления весовых коэффициентов признаков (ВКП) z_k , $k = \overline{1, m}$ и весовых коэффициентов тестов (ВКТ) формировать совокупности всех различных пар объектов (g -эталонов) из разных образов для каждого признака z_k , в виде мультимножества P_k (допускающего кратность элементов, см. подробно пример в приложении 4).

7: Вычислить значения ВКП и ВКТ на основе одного из методов (гл. 4, пп.4.5, 4.6).

8: Распознаваемый объект Y сравнивается с каждым обучающим объектом Y_i по каждому тесту (шаг 3). Объект Y получает голос за принадлежность классу ν , если $Y_i \in \nu$ и описания объектов Y и Y_i совпадают по всему множеству (части множества) признаков τ (в этом случае $B(Y, Y_i, \tau) = 1$). Для каждого класса

$\cup \subset (\Omega_1, \dots, \Omega_l)$ вычисляется оценка принадлежности $G(Y, \nu)$ объекта Y классу ν , которая имеет вид:

$$G(Y, \nu) = \frac{1}{|V_\nu|} \sum_{Y_l \in \nu} \sum_{\tau \in T} \gamma_l \cdot w_\tau \cdot B(Y, Y_l, \tau), \quad (4.8)$$

где $V_\nu = \nu \cap \{Y_1, \dots, Y_m\}$, $|X|$ – мощность множества X , γ_l – параметр, характеризующий представительность объекта Y_l ; w_τ – параметр, характеризующий представительность теста τ , $\tau \in \{T_j, j = \overline{1, N_T}\}$, N_T – количество найденных тестов.

Вернуть: эталонное описание $Q' = \|q'_{jk}\|_{n' \times m}$, $R' = \|r'_{jl}\|_{n' \times l}$; множество решающих правил T ; итоговое решающее правило; номер класса распознаваемого объекта.

Замечание. Традиционно в тестовом распознавании итоговое решение принимается на основе голосования на множестве тестов [55] по оценкам принадлежности $G(Y, \nu)$ (4.8). Объект Y относится к тому классу, который имеет наибольшую оценку. Преимущество применения модифицированного МАИ для этих целей заключается в системном эффекте данного метода, а именно, учете взаимозависимости (взаимовлияния) тестовых наборов признаков. В п. 3.5 (теорема 3.1) показано, что ММАИ приводит к обычному голосованию, если сравниваемые наборы независимы.

4.4. Алгоритм распознавания состояния ДО на основе формализма мультимножеств

Пусть задана система градаций квантования (СГК) $G = \{\kappa_\nu, \nu : \nu \in U\}$, $U = \{\nu_j : j = \overline{1, N_U}\}$. Поставим в соответствие каждому выделенному состоянию ДО Ω_i мультимножество $\Theta_i = \{k_{\Theta_i}(x) \bullet x | x \in U, k_{\Theta_i}(x) \in Z_+\}$ (опр. 1.15). Напомним, что любое мультимножество $G = (g_1, g_2, \dots, g_n)$ из n разных элементов можно задать его основанием $S(G) = (x_1, x_2, \dots, x_n)$ – кортежем всех разных элементов из U и кортежем кратностей $[G] = (k_1, k_2, \dots, k_n)$ элементов основания (первичная спецификация мультимножества). Поскольку Θ_i содержит элементы, являющиеся случайными числами, то Θ_i – стохастическое мультимножество (СММ).

Алгоритм 4.2. Идентификация состояний ДО на основе формализма мультимножеств

Вход: обучающая выборка – размеченный временной ряд (один или несколько); результат новых наблюдений $y(t, k) = (y(t-k), y(t-k+1), \dots, y(t))$.

Выход: метка (номер) состояния ДО для наблюдений $y(t, k)$.

1: Сопоставляем каждому состоянию ДО $\Omega_i, i = \overline{1, I}$ СММ $\Theta_i = \{k_{v_1}^i \bullet v_1, k_{v_2}^i \bullet v_2, \dots, k_{v_{N_i}}^i \bullet v_{N_i}\}$, сформированное по обучающей выборке. Мультимножествам Θ_i будут соответствовать основания $S(\Theta_i) = \{v_1, v_2, \dots, v_{N_i}\}$ и первичные спецификации $[\Theta_i] = \{k_{v_1}^i, k_{v_2}^i, \dots, k_{v_{N_i}}^i\}$. Элементами основания могут быть или значения, полученные традиционным логарифмированием $v_l = \lceil \log(y_j) \rceil, j \in T_i$, где $\lceil x \rceil$ - целая часть числа x , T_i - интервал временных отсчетов, соотнесенный с Ω_i , или g-эталон (алгоритм 2.1).

2: Определяем оценки распределения вероятностей $\{p_v^i, v \in S(\Theta_i)\}$ уровней значений, характерных для каждого состояния ДО по формуле
$$p_{v_j}^i = k_{v_j}^i \left(\sum_{l \in S(\Theta_i)} k_l^i \right)^{-1}, j = \overline{1, N_i}, i = \overline{1, I}.$$

3: По результатам наблюдений $y(t, k) = (y(t-k), y(t-k+1), \dots, y(t))$ (k -окну) формируем СММ $\Theta = \{k_{s_1} \bullet s_1, k_{s_2} \bullet s_2, \dots, k_{s_{N_{tk}}} \bullet s_{N_{tk}}\}$, где N_{tk} - число различных значений $s_l = \lceil \log(y_j) \rceil, j \in T(t, k)$, где $T(t, k)$ - интервал временных отсчетов, соотнесенный с k -окном.

4: Вычисляем информационные расстояния между СММ $d(\Theta_i, \Theta), i = \overline{1, I}$ по $J \geq 1$ выбранным мерам сходства/различия (4.4)-(4.7), (4.11), (5.7).

5: Решение принимаем в пользу состояния ДО i_0 , имеющего наибольший весовой коэффициент по алгоритму 2.3, гарантирующему парето-оптимальность многокритериального выбора.

Вернуть: i_0 .

4.5. Методы оценивания весовых коэффициентов признаков в тестовом распознавании образов

Одной из наиболее важных проблем при создании интеллектуальных систем выявления закономерностей и принятия решений является проблема анализа признакового пространства на предмет выделения наиболее значимых признаков и оценивания величины их значимости. Несмотря на то, что этой проблеме посвящено большое количество публикаций, (например, обзор в [27]), до настоящего времени отсутствуют общие рекомендации по определению весовых коэффициентов признаков (ВКП) и весовых коэффициентов тестов (ВКТ).

В интеллектуальных системах, основанных на методах тестового распознавания образов, для принятия решений используются «хорошие» тесты, то

есть тесты, содержащие меньшее количество признаков и с большим весом, где под весом теста традиционно понимается сумма ВКП.

Известно сравнительное исследование [27, 132] практически всех существующих на данный момент методов определения коэффициентов важности критериев на основе понятий современной теории измерений. Однако, непосредственное применение «критериальных» методов для оценки значимости (весовых коэффициентов) признаков не представляется возможным в силу специфики рассматриваемых задач. Так, используемое в задачах тестового распознавания понятие признака и критерия (в многокритериальной задаче) не являются равнозначными. Критерии, как правило, представляют собой качественные характеристики, которые для определения весовых коэффициентов признаков искусственно трансформируются экспертами в количественные оценки, причем для каждого критерия может быть выбрана своя оценочная шкала. Следует отметить, что более детально представлены результаты по многокритериальным функциям предпочтения, использующих базовые шкалы оценок альтернатив по критериям [132]. Функция предпочтения, как правило, представляет собой отображение множества альтернатив на числовую ось, и лучшей альтернативе приписывается большее число.

Постановка задачи. Пусть по матрицам \mathbf{Q} и \mathbf{R} построены все (часть) тупиковые тесты, представленные матрицей тестов \mathbf{T} , строки которой сопоставлены тестам, а столбцы – характеристическим признакам, и определено число различающих пар «образ-образ» («объект-образ», «объект-объект») по каждому характеристическому признаку.

Требуется определить весовые коэффициенты характеристических признаков, входящих в объединение всех (части) тупиковых тестов.

При этом не исключается возможность достаточно большой размерности признакового пространства и наличия взаимозависимости признаков.

Решение задачи. Кратко изложим существующие и новые методы определения весовых коэффициентов признаков и тестов, а также введем эмпирический функционал качества, определяющий среднюю ошибку распознавания на контрольной выборке по каждому из рассмотренных методов. На основе эмпирического функционала качества будет построена коллективная оценка признаков (тестов), при этом вклад разных методов будет осуществлен пропорционально эффективности соответствующего метода на контрольной выборке.

4.6. Дискретный и статистический подходы к оцениванию весовых коэффициентов признаков и тестов

Обозначим через w_m^l – весовой коэффициент признака, соответствующего m -му столбцу матрицы \mathbf{Q} ($m = \overline{1, M}$), по l -му методу.

Метод 1 (ВКП на основе информационного веса). Предложенная Ю.И. Журавлевым в [55] для тестовых алгоритмов распознавания образов мера важности признака (информационный вес), как отношение числа вхождений признака во все безызбыточные (тупиковые) тесты к числу всех безызбыточных тестов, логична, интуитивно прозрачна и легко интерпретируема. Эта мера важности пригодна для задач распознавания образов малой размерности, так как количество безызбыточных тестов при числе признаков, равном несколько десятков, может быть весьма большим [52] и нахождение всех безызбыточных тестов становится нереальным при применении даже современных компьютерных технологий.

При большой размерности задачи данный метод уступает другим, поскольку чем меньше количество безызбыточных тестов будет использовано для вычисления информационного веса признаков, тем менее точна оценка ВКП.

Метод 2, основанный на различающей способности признака, предложенный А.Е. Янковской [175], весьма прозрачен, прост и эффективен при репрезентативной выборке. Весовой коэффициент w_m^2 m -го признака по этому методу определяется по формуле:

$$w_m^2 = \frac{\sum_{r=1}^{K-1} \sum_{t=r+1}^K \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} \delta_{ij}^m}{\sum_{i=1}^{K-1} \sum_{j=i+1}^K \sigma_i \sigma_j}, \quad (4.9)$$

где K – число выделенных образов; N_f – число строк в описании f -го образа ($f \in \{r, t\}$), $\eta_{ij}^m = 0$, если $q_{im}=q_{jm}=0$ или $q_{im}=q_{jm}=1$ (q_{im} и q_{jm} – элементы матрицы \mathbf{Q} из разных образов); $\eta_{ij}^m = x_i x_j 2^{d_i^- + d_j^-}$ (d_i^- – число значений «-» в i -ой строке матрицы \mathbf{Q} , x_i – коэффициент повторения i -ой строки), если $q_{im}=0$ и $q_{jm}=1$ или $q_{im}=1$ и $q_{jm}=0$; $\eta_{ij}^m = x_i x_j 2^{d_i^- + d_j^- - 1}$, если $q_{im}=\langle\langle-\rangle\rangle$ и (или) $q_{jm}=\langle\langle-\rangle\rangle$; σ_j – число объектов в j -ом образе ($j = \overline{1, K}$), вычисляемое по формуле: $\sigma_j = \sum_{k=1}^{N_j} x_k 2^{d_k^-}$; $m \in \{1, 2, \dots, M\}$.

Данный метод апробирован более чем на 70-ти базах данных и знаний, например (см. библиографию в [99]), при этом количество объектов доходило до нескольких тысяч, а признаков – до нескольких сотен. При большой размерности решаемых задач второй метод более эффективен (по качеству решения (точности оценки ВКП), стоимости и временным затратам) по сравнению с первым методом.

Метод 3, основанный на нахождении математического ожидания весовых коэффициентов признаков, изложенный И.Ф. Шахновым в [162] для оценки целей многоцелевой планируемой операции, адаптирован авторами [171] к вычислению весовых коэффициентов признаков в интеллектуальных системах, основанных на тестовом распознавании образов.

Метод основан на выборе в качестве «точечных оценок» признаков их значений в центре тяжести области допустимых значений U , то есть математическое ожидание значений $z_m, m = \overline{1, M}$, в случае равномерного распределения вектора $Z = (z_1, z_2, \dots, z_M)$ в U . При этом используются следующие предположения: а) множество входящих в тест признаков упорядочены и перенумерованы в порядке убывания их важности (установление предпочтения отведено лицу, принимающему решение); б) известны результаты сравнения с точки зрения их предпочтительности отдельного признака z_m и «комплексного» (следуя терминологии [162]), представляющего собой объединение двух следующих $z_{m+1}, z_{m+2} : z'_m = z_{m+1} \cup z_{m+2}, m = \overline{1, M-2}$, то есть для каждой пары (z'_m, z_m) имеет место одно из следующих трех соотношений:

$$z'_m \succ z_m, z_m \sim z'_m, z_m \succ z'_m, m = \overline{1, M-2}$$

Метод 3 включает 2 этапа: 1) нахождение коэффициентов признаков z_{M-2}, z_{M-1}, z_M , обозначенных через $\bar{z}_{M-2}, \bar{z}_{M-1}, \bar{z}_M$; 2) нахождение коэффициента \bar{z}_m признака z_m с учетом заданного предпочтения относительно признака z_m и $z'_m = z_{m+1} \cup z_{m+2}, m = \overline{1, M-3}$.

При выполнении вышеупомянутых условий метод 3 дает возможность прогнозировать характер изменения количественных значений ВКП, двигаясь от более предпочтительного к менее предпочтительному, а также удобен при минимальной информации (не обязательно числовой) о признаках при использовании интервальной шкалы или шкалы порядка.

При использовании шкалы пропорциональных оценок (наиболее предпочтительной в принятии решений) метод 3 менее точен, т.к. при одинаково установленном предпочтении будет выдавать одинаковые оценки весов, несмотря на существенные количественные различия между ними. Данный метод можно также рассматривать как возможный способ (при минимальной информации о признаках) перевода заданной ранговой шкалы в количественную шкалу отношений.

Пример 4.1. Пусть задана двоичная матрица описаний объектов (\mathbf{Q}) в пространстве 4-х характеристических признаков и одностолбцовая матрица различий (\mathbf{R}), строки которой сопоставлены строкам матрицы \mathbf{Q} , а столбец – классификационному признаку, разбивающему объекты на классы эквивалентности (6 образов).

$$\mathbf{Q} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} ; \quad \mathbf{R} = \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{matrix} \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}.$$

Полученные методом 3 весовые коэффициенты признаков, равны 0.56; 0.6; 0.6; 0.48. а по методу 2 0.225; 0.376; 0.274; 0.125 соответственно.

Результаты сравнения двух методов получения весовых коэффициентов признаков говорят о безусловном преимуществе по точности вычисления коэффициентов по методу 2 по сравнению с методом 3, поскольку с какими бы значениями весовых коэффициентов характеристических признаков с теми же предпочтениями для различных матриц \mathbf{Q} и \mathbf{R} с одним и тем же числом столбцов матрицы \mathbf{Q} мы не брали, получим по методу 3 одинаковые значения весовых коэффициентов признаков.

При выполнении вышеупомянутых условий метод 3 дает возможность прогнозировать характер изменения количественных значений ВКП, двигаясь от более предпочтительного к менее предпочтительному, а также удобен при минимальной информации (не обязательно числовой) о признаках при использовании интервальной шкалы или шкалы порядка.

При использовании шкалы пропорциональных оценок (наиболее предпочтительной в принятии решений) метод 3 менее точен, т.к. при одинаково установленном предпочтении будет выдавать одинаковые оценки весов, несмотря на существенные количественные различия между ними. Данный метод можно также рассматривать как возможный способ (при минимальной информации о признаках) перевода заданной ранговой шкалы в количественную шкалу отношений.

Метод 4 на основе дивергенции Кульбака [112, 154]. Данный метод относится к статистическим и вычисление весовых коэффициентов признаков осуществляется на основе дивергенции Кульбака, которая определяет предельно достижимую вероятность ошибки при распознавании образов и позволяет ранжировать признаки по их различающей способности: наибольшей различающей способностью (информативностью) обладает тот признак, для которого дивергенция информации максимальна. При этом, возможна сравнительная оценка различающей способности двух признаков по отношению их дивергенций информации. Кратко изложим метод 4.

Пусть имеется некоторый признак A , который может принимать градации A_1, A_2, \dots, A_k . Пусть для первого образа H_1 эти градации появляются с вероятностями p_1, p_2, \dots, p_k , а для второго образа H_2 – с вероятностями q_1, q_2, \dots, q_k . Тогда дивергенция Кульбака имеет вид [112]:

$$I(H_1 : H_2) = \sum_{i=1}^k (p_i - q_i) \cdot \ln \frac{p_i}{q_i}. \text{ Эта величина определяет предельно достижимую}$$

вероятность ошибки при распознавании образов H_1 и H_2 .

Вероятности p_1, p_2, \dots, p_k и q_1, q_2, \dots, q_k оцениваются по обучающей выборке: $\hat{p}_i = n_i/N, \hat{q}_i = m_i/M$, где N, M – число образов H_1 и H_2 , и в них признак A в i -й градации встретился n_i и m_i раз ($i = \overline{1, k}$), соответственно.

Тогда формула оценки дивергенции информации по Кульбаку имеет вид:

$$\hat{I}(H_1 : H_2) = \sum_{i=1}^k (\hat{p}_i - \hat{q}_i) \cdot \ln \frac{\hat{p}_i}{\hat{q}_i} = \sum_{i=1}^k \left(\frac{n_i}{N} - \frac{m_i}{M} \right) \cdot \ln \frac{n_i M}{m_i N}.$$

Показано в [112], что для N и M много больше 1 оценка дивергенции информации является асимптотически несмещённой. Учитывая независимость образов в обучающих выборках, показана сходимость оценки \hat{I} к I в средне квадратичном смысле.

Метод вычисления ВКП на основе дивергенции Кульбака привлекателен тем, что позволяет не только ранжировать признаки, но и количественно их сравнивать для задач большой размерности.

Метод 5, основанный на частотах встречаемости признаков для каждого из классов (образов) [99]. Основу данного метода составляет заданная информация о частотах встречаемости признаков (оценки вероятностей проявления признака) для каждого из классов (образов). Для ряда проблемных областей, например, в медицине для некоторых заболеваний на основе статистического анализа известны вероятности проявлений тех или иных симптомов. При условии, что по каждому признаку заданы частоты попадания объектов в определенные образы (классы) или вероятности единичных значений признаков для каждого образа, предлагается следующая формула вычисления весового коэффициента m -го признака w_m^5 ($m = \overline{1, M}$):

$$w_m^5 = \left(C_K^2 \right)^{-1} \sum_{s=1}^{K-1} \sum_{j=s+1}^K \left(p_{ms}(1 - p_{mj}) + p_{mj}(1 - p_{ms}) \right), \quad (4.10)$$

где p_{ik} - вероятность единичного значения i -го признака по k -му образу ($k \in \{s, j\}$, $s = \overline{1, K-1}$, $j = \overline{s+1, K}$), K – количество образов, C_K^2 – число сочетаний из K по 2. Данная формула справедлива, если выборка репрезентативна и объемы выборок по каждому образу совпадают.

Перечисленные выше методы и существующие методы принятия решения либо требуют выполнения условия независимости признаков (критериев) по предпочтению, либо просто не учитывают реально возможную их взаимозависимость, что может привести к неточности при принятии решений в интеллектуальной тестовой распознающей системе.

Оригинальные методы вычисления ВКП и весовых коэффициентов тестов (ВКТ) далее существенно будут опираться на формализм мультимножеств и модификацию МАИ (гл. 3).

4.7. Метод оценивания весовых коэффициентов признаков и тестов на основе формализма мультимножеств и нелинейной модификации метода анализа иерархий

Основанный на формализме мультимножеств [131] метод определения весовых коэффициентов признаков (далее на него будем ссылаться как на *метод б*) при принятии решения в интеллектуальных тестовых распознающих системах с матричным представлением данных и знаний учитывает вклад признаков в распознающую способность теста с учетом их взаимозависимости и базируется на представлении совокупности всех различимых пар объектов из разных образов для каждого признака $z_m, m = \overline{1, M}$, в виде мультимножества, применении модификации метода анализа иерархий Саати, использующего парные сравнения признаков на основе специальным образом выбранной меры относительной важности признаков, учитывающей их взаимозависимость.

Соответствующий признаку z_m m -й столбец матрицы \mathbf{Q} порождает совокупность P_m различимых этим признаком пар объектов из разных образов $\{(i-j) \mid i \in F_r, j \in F_t, r \neq t, q_{im} \neq q_{jm}\}$, где $F_l = \{f_l, f_l+1, \dots, f_l+N_l-1\}$, N_l – число строк в l -ом образе, f_l – номер 1-й строки в матрице \mathbf{Q} для l -го образа, $l \in \{r, t\}$. Отметим, что количество пар $(i-j)$ данной совокупности вычисляется по следующей формуле: $\eta_{ij}^m = x_i x_j 2^{d_i} 2^{d_j}$

Поставим данному признаку z_m в соответствие совокупность различимых m -м признаком пар объектов из разных образов и будем говорить, что признак z_m порождает мультимножество $P_m = \{k_{P_m}(u) \bullet u \mid u \in U, k_{P_m}(u) \in Z_+\}$.

Таким образом представленные признаки являются множествами P_m с повторяющимися элементами $u \in U$ (мультимножествами), при этом, мощность мультимножества $|P_m| = \sum_{P_m} k_{P_m}(u)$ определяется как общее число его элементов; размерность мультимножества $/P_m/ = \sum_{P_m} \chi_{P_m}(u)$ – как общее количество различных элементов (функция $\chi_{P_m}(u) = 1$ при $u \in P_m$, и $\chi_{P_m}(u) = 0$ при

$u \notin P_m$). Если $k_{P_m}(u) = \chi_{P_m}(u)$, то мультимножество P_m является обычным множеством.

Метод состоит из трех этапов, на каждом из которых формируется матрица парных сравнений (МПС) $\mathbf{A} = \left\| a_{ij} \right\|_{g \times g}$ признаков $(z_{i_1}, z_{i_2}, \dots, z_{i_h}, \dots, z_{i_g})$ ($i_h \in \{1, \dots, M\}$, g – количество признаков в тесте) на основе определенной меры относительной важности признака i над признаком j , в качестве которой поэтапно выбираются величины (верхний индекс элемента МПС a_{ij}^s обозначает номер этапа, $s \in \{1, 2, 3\}$):

$$a_{ij}^1 = \frac{|P_i|}{|P_j|}; a_{ij}^2 = \frac{\delta(|P_i - P_j|)}{\delta(|P_j - P_i|)}; a_{ij}^3 = \frac{\delta(|P_i - P_j|)}{\delta(|P_j - P_i|)}, \quad (4.11)$$

где $|P_i|$, $|P_j|$ – мощность и размерность i -го мультимножества, сопоставленного признаку z_i , соответственно, $P_i - P_j$ – разность мультимножеств, соответствующих признакам z_i и z_j , где функция $\delta(x) = x$, если $x \neq 0$, и $\delta(x) = 1$, иначе.

На s -м этапе по 6-му методу вычисляется g – компонентный вектор значений весовых коэффициентов признаков – $\mathbf{W}_s^6 = (w_{si_1}^6, w_{si_2}^6, \dots, w_{si_g}^6)$, совпадающий со значением нормализованной оценки (как среднее геометрическое элементов строки [141]) главного собственного вектора МПС (по Саати – локальные приоритеты). Вектор $\mathbf{W}^6 = (\mathbf{W}_1^6 \cdot \mathbf{W}_2^6 \cdot \mathbf{W}_3^6)^{1/3}$ представляет собой обобщенные значения весовых коэффициентов признаков (глобальные приоритеты), входящих в тест.

Вышеуказанные меры позволяют учесть не только общие свойства сравниваемых признаков (показатели (индексы) сходства или различия), но, что особенно важно, и уникальные их свойства (степень приоритетности одного признака над другими), в отличие от метода 3.

Отметим, если знания, представленные матрицами \mathbf{Q} , \mathbf{R} , репрезентативны, а признаки статистически независимы, то результаты, полученные по методам 2, 5 и первому этапу метода 6, совпадают.

Установлена методологически и практически важная связь между формулой (4.9), дающей значения весовых коэффициентов признаков на основе их разделяющей способности и оценкой главного собственного вектора, полученной на первом этапе метода 6 (теорема 4.1).

Теорема 4.1. Нормализованная оценка главного собственного вектора матрицы парных сравнений признаков $\mathbf{W}_1^6 = (w_{1,i_1}^6, w_{1,i_2}^6, \dots, w_{1,i_g}^6)$ с мерой относитель-

ной важности признака i над признаком j , равной $a_{ij}^1 = \frac{|P_i|}{|P_j|}$, совпадает с нормализованным вектором $W^2 = (w_{i_1}^2, w_{i_2}^2, \dots, w_{i_g}^2)$, $i_m \in \{1, \dots, M\}$, полученным по формуле (4.9).

4.8. Метод оценивания весовых коэффициентов признаков и тестов при большой размерности признакового пространства

Известно достаточно большое число методов снижения размерности признакового пространства, в частности, можно упомянуть регрессионный, дискриминантный и кластерный анализы [25]. Однако такие процедуры требуют репрезентативных и статистически однородных данных, которые могут отсутствовать при решении новых задач. Как правило, сбор подобной статистики может потребовать больших временных затрат, что не всегда представляется возможным.

Среди способов снижения размерности признакового пространства важно отметить методы векторной стратификации [27], основанные на процедурах построения структурированного многокритериального пространства и разбиения его на заданное число упорядоченных слоев (страт) и теорию гранулирования информации Л.А. Заде.²⁴, основанную на использовании аппарата нечетких множеств и лингвистических переменных.

Рассмотрим метод оценивания ВКП, частично решающий проблему нахождения ВКП для случая, когда размерность признакового пространства достаточно велика (метод 7). Как известно, данная задача до сих пор является нетривиальной не только в случае задания предпочтений экспертами вручную, но и в случае, когда относительная важность ВКП вычисляется по предложенным выше формулам (методы 1-6). Например, для метода 6 в случае 50-ти признаков, используемых в задачах диагностики, количество парных сравнений составит 1225, что является достаточно трудоемким процессом.

Метод 7 существенно опирается на идею В.Д. Ногина [126] о «базисных» элементах (расположенных выше (либо ниже) главной диагонали), на основе которых затем находится искомый весовой вектор (просто и без ошибок вычислительного характера). При этом выбор конкретного «базисного» набора соответствует той или иной схеме сравнения объектов, которую можно выбирать с целью получения наиболее надежных результатов. Метод 7, во-первых, основан на совместной матрице A и, таким образом, избавлен от «модельной» ошибки;

²⁴ Zadeh L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic // Fuzzy Sets and Systems. -1997. - September. - Vol. 90. - No. 2. - Pp. 111-127.

во-вторых, он обеспечивает количество требуемых сравнений существенно меньшее традиционно требуемых.

Подход к решению данной проблемы, связанный с упрощенным вариантом МАИ, авторами адаптирован следующим образом.

Сначала выбираются $M-1$ базисных элементов, на основе которых определяются все остальные элементы МПС. Способ выбора базисных элементов является задачей творческой и зависит от модели задачи (и решения эксперта), в частности, от выбранной меры сравнения относительной важности признаков. Применительно к данной задаче рассмотрим два следующих варианта сокращения сравнений в рамках указанного подхода.

Первый вариант связан с выбором некоторого «идеального» признака-образца, в качестве которого может служить либо признак с наибольшим значением ВКП, вычисленным по формуле (4.9), либо признак, которому будет сопоставлено объединение $(\bigcup_m P_m)$ всех мультимножеств, порожденных соответствующими признаками z_m ($m = \overline{1, M}$).

Теорема 4.2. Пусть МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$ построена на основе одной из мер относительной важности (4.11) и в соответствии с формулой:

$$a_{ij} = a_{i1} \cdot a_{1j} = \frac{a_{1j}}{a_{1i}}, \quad i = \overline{2, M}, \quad j = \overline{1, M}. \quad (4.12)$$

Тогда, во-первых, существует единственная МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$, обладающая всеми свойствами (3.1); во-вторых, компоненты весового ненормализованного вектора $W_1^7 = (w_{11}^7, w_{12}^7, \dots, w_{1M}^7)^T$ вычисляются по формуле:

$$w_{1i}^7 = \frac{a_{1M}}{a_{1i}}, \quad i = \overline{1, M}. \quad (4.13)$$

Здесь первый нижний индекс означает номер способа выбора базисных элементов в рамках 7-го метода.

Таким образом, компоненты вектора ВКП составляют последний столбец МПС A , элементы которой получены по формулам (4.12) и (4.13).

Преимущество данного способа с точки зрения трудоемкости вычислений при $M \geq 4$ очевидно: вместо вычисления $M \cdot (M-1)/2$ величин потребуется вычислить только $2 \cdot M - 3$ величин, то есть более чем в $M/4$ раза меньше. Действительно, рассмотрев отношение сравниваемых величин, получим выражение, из которого следует справедливость данного вывода:

$$\frac{M \cdot (M-1)/2}{2 \cdot M - 3} = \frac{1}{4} \cdot \left(M - 1 + \frac{3 \cdot (M-1)}{2 \cdot M - 3} \right) = \frac{M}{4} \cdot \left(1 + \frac{1}{2 \cdot M - 3} \right) > \frac{M}{4}, \quad M > 3.$$

К тому же, в случае экспертного задания оценок, формула (4.12) обеспечивает не только экономию времени, но и согласованность МПС, то есть избавляет от «модельной» ошибки.

Второй вариант предполагает следующий набор базисных элементов: $a_{12}, a_{23}, \dots, a_{M-1,M}$, который реализует схему последовательного сравнения (некоторый «идеальный образец» сравнивается с 1-м, затем 1-й сравнивается со вторым, и т.д.). Сформулируем результат.

Теорема 4.3. Пусть МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$ построена по следующим правилам: 1) элементы $a_{12}, a_{23}, \dots, a_{M-1,M}$ являются базисными и определяются по формуле:

$$a_{ij} = \frac{\delta(|P_i - P_j|)}{\delta(|P_j - P_i|)} \left(a_{ij} = \frac{\delta(/P_i - P_j/)}{\delta(/P_j - P_i/)} \right), \quad i = \overline{1, M-1}, j = \overline{2, M}; \quad (4.14)$$

2) остальные элементы определяются по формуле:

$$a_{ij} = a_{i,j-1} \cdot a_{j-1,j}, \quad i = \overline{1, M-2}, i < j-1. \quad (4.15)$$

Тогда МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$ удовлетворяет всем требованиям относительных весов (3.1), и компоненты ненормированного собственного вектора W_2^7 определяются по формуле:

$$w_{2i}^7 = a_{i,i+1} \cdot a_{i+1,i+2} \cdots a_{M-1,M}, \quad i = \overline{1, M-1}, w_M = 1. \quad (4.16)$$

Отметим, что второй способ предполагает вычисление компонентов собственного вектора по базисным элементам, определяемым по формулам (4.14), (4.15), и количество требуемых сравнений определится числом $M-1$, что в $M/2$ раз меньше традиционно требуемых сравнений $M(M-1)/2$.

4.8.1. Сравнительный пример определения ВКП по разным методам

Пусть заданы матрицы **Q**, **R**, **T**, представленные на рис. 4.1 (для простоты изложения матрица различий **R** представлена одним столбцом). Вычислим для одного из тестов значения ВКП (табл. 4.1) по методам 1–7, например, для $\tau = (z_5, z_9, z_{11})$. Отметим, что общее количество тестов для рассматриваемого примера равно 6. Для удобства соотнесения номеров признаков в матрицах **Q** и **T**, в матрице **T** использованы все 13 признаков, хотя столбцы 1, 2, 8, 10 избыточны, поскольку соответствующие им признаки не входят во множество безызбыточных тестов, представленных матрицей **T**.

Так как каждая из строк 3, 6, 7 представляет собой интервал булева пространства (содержит по одному символу «–») и представляется двумя строками, то число строк матрицы **Q** увеличится на 3.

В приложении 4 содержится построение оценки ВКП по методу 6 как иллюстрация использования формализма мультимножеств.

Таблица 4.1. Значения весовых коэффициентов признаков

№	Весовые коэффициенты признаков	z_5	z_9	z_{11}
1	W^1	0.462	0.077	0.462
2	W^2	0.326	0.349	0.326
3	W^3	0.361	0.444	0.194
4	W_1^6	0.326	0.349	0.326
5	W_2^6	0.316	0.375	0.308
6	W_3^6	0.385	0.308	0.308
7	W^6	0.341	0.343	0.314
8	W^7	0.325	0.375	0.300

Строки 1 – 3 (W^1 , W^2 , W^3) табл. 4.1 содержат значения ВКП, полученных по первым трем методам. Отметим, что по методу 3 получим одинаковые значения ВКП для различных матриц **Q** и **R** с одним и тем же числом столбцов матрицы **Q**, если исходные предпочтения между признаками одинаковые.

Строки 4 – 5 табл. 4.1 содержат значения ВКП, полученных на каждом из этапов по методу 6, и представляют самостоятельный интерес. Заметим, что оценки ВКП, полученных по методу 2 и 1-му этапу метода 6 совпадают, а также имеет место изменение приоритетов признаков, весовые коэффициенты которых вычисляются по методам 2 и 6 (7-я строка), что объясняется более подроб-

ным учетом информации по методу 6, использующему парные сравнения признаков на основе 3-х различных мер (4.11).

Строка 8 (метод 7) содержат значения ВКП, учитывающие взаимозависимость признаков, и имеющие заметное преимущество для решения реальных задач большой размерности.

Сравним численные значения весовых коэффициентов признаков, полученных по методу 6 и методу 7, использующему упрощенный МАИ. Ввиду громоздкости МПС при большой ее размерности, а также имея в виду, что интерес, главным образом, представляет а) абсолютное отклонение между оценками ВКП по данным методам, то есть $|W_2^6 - W_2^7|$, где W_2^6, W_2^7 – нормализованный собственный вектор МПС на этапе 2 метода 6, W_2^7 – нормализованный собственный вектор МПС по методу 7 с использованием 2-го способа выбора базисных элементов; б) факт совместности МПС или значение индекса согласованности в указанных методах $|a_{ij} - \frac{w_i}{w_j}|$, $i, j = \overline{1, M}$, рассмотрим получение значений весовых коэффициентов признаков для вышеупомянутого теста $\tau = (z_5, z_9, z_{11})$, построенного по матрицам Q, R, T, представленных на рис. 4.1. В МПС признаков на основе 2-й меры относительной важности (4.11) расположим признаки по невозрастанию мощностей соответствующих разностей мультимножеств (табл. 4.2), учитывая, что $|P_5 - P_9|=13$, $|P_9 - P_5|=15$, $|P_5 - P_{11}|=14$, $|P_{11} - P_5|=14$, $|P_9 - P_{11}|=10$, $|P_{11} - P_9|=8$.

Соответствующая МПС признаков на основе указанной меры относительной важности и вектор ВКП W_2^6 представлены в табл. 4.2.

Таблица 4.2. Весовые коэффициенты признаков по методу 6

	z_5	z_9	z_{11}	W_2^6
z_5	1	15/13	5/4	0.375
z_9	13/15	1	14/14	0.316
z_{11}	4/5	14/14	1	0.308

МПС признаков на основе выбранной меры относительной важности и упрощенного МАИ (формулы (3.6), (3.7)) и вектор ВКП W_2^7 представлены в табл. 4.3.

Таблица 4.3. Весовые коэффициенты признаков по методу 7

	z_5	z_9	z_{11}	W_2^7
z_5	1	15/13	5/4	0.375
z_9	13/15	1	(5·13)/(4·15)	0.325
z_{11}	4/5	(4·15)/(5·13)	1	0.300

Таким образом, относительное отклонение между оценками ВКП по данным методам $|W_2^6 - W_2^7| = (0.000, 0.009, 0.008)^T$.

Обратим также внимание на погрешность, связанную с индексом совместности МПС по методу 6, основанному на мультимножествах и МАИ, имея в виду, что уровень отношения согласованности равен 0.001 (для применимости МАИ для данной МПС желательным считается значение индекса согласованности, меньшее 0.1 [141]):

$$\left\| a_{ij} - \frac{w_i}{w_j} \right\| = \begin{bmatrix} 0 & 0.031 & 0.033 \\ 0.023 & 0 & 0.027 \\ 0.022 & 0.026 & 0 \end{bmatrix}.$$

Таким образом, «модельная ошибка» в методе 6, основанном на мультимножествах и формулах (4.11), оказалась больше (максимальное отклонение равно 0.033), чем максимальное отклонение $|W_2^6 - W_2^7|$, равное 0.008. При этом выигрыш по методу 7, позволяющему в 2 раза сократить количество сравнений и упростить процедуру вычисления собственного вектора, существенен при большом числе признаков.

4.9. Показатели эффективности методов оценивания признаков, тестов и принятие итогового решения

Совместное использование в едином решающем правиле разнотипных моделей является перспективным способом наиболее полного учёта априорной информации.

Каждый из рассмотренных в гл. 4 методов оценивания ВКП, входящих в тест, тем самым порождает коэффициенты сходства S_k (k – номер образа, $k = \overline{1, K}$) объектов внутри образа (внутриклассового сходства) и коэффициенты сходства S_k^x исследуемого (распознаваемого) x -го объекта с k -м образом (классом), формулы для вычисления которых при $q_{ij} \in \{0, 1, \llcorner\}$ для 2-го и 5-го методов даны в [23, 175] и не приводятся в виду их громоздкости. Отношение G_k коэффициентов S_k^x к коэффициентам S_k (по каждому образу отдельно) вида $G_k = S_k^x / S_k$ будем называть условной степенью близости распознаваемого (исследуемого) объекта к k -му образу.

Принятие решений относительно исследуемого объекта по каждому тесту в рамках логико-комбинаторного подхода осуществляется с использованием порогового значения условной степени близости данного объекта к образам (классам).

Принятие итогового решения осуществляется на основе специальной процедуры голосования по всем решающим правилам, построенным по соответствующим тестам.

Для методов $l = \overline{1, 7}$, $l \neq 2, 5$ предлагается скорректировать эти формулы, внося в них поправочные коэффициенты в виде значений отношений весовых коэффициентов w_m^l / w_m^2 , $l = \overline{1, 7}$, $l \neq 2, 5$, $m = \overline{1, M}$.

Поставим задачу ранжирования тестов по важности, которая тесно связана с задачей нахождения «лучшего» метода определения весов тестов. Представляет интерес и метод, позволяющий построить коллективную оценку для тестов.

В качестве показателя эффективности метода получения ВКП и, соответственно, тестов рассмотрим функционал качества, зависящий от доли ошибок, которые допускает алгоритм на рассматриваемой выборке, что считается традиционным в теории распознавания образов. При этом наилучшим для рассматриваемой задачи считается алгоритм, допускающий на контрольной выборке минимум ошибок.

Обозначим через y_{ij}^r – номер образа, которому принадлежит j -й контрольный объект по l -му методу и r -му тесту ($l = \overline{1, L}$, $r = \overline{1, N_T}$, $j = \overline{1, N}$) и введем функционал качества, определяющий среднюю ошибку распознавания по l -му методу (отнесение j -го объекта к y_{ij}^r -му образу) как показатель эффективности r -го теста по l -му методу:

$$\Phi_l^r = \frac{1}{N} \sum_{j=1}^N \gamma_{ij}^r \quad (4.17)$$

где $\gamma_{ij}^r = (G_k - 1)^2$, если j -й контрольный объект принадлежит k -му образу и $\gamma_{ij}^r = 1$, иначе; величина G_k вычисляется по вышеприведенной формуле при $k = y_{ij}^r$, где j – номер исследуемого объекта x .

Метод l ($l = \overline{1, L}$) определения ВКП, доставляющий минимум этому функционалу, будет «лучшим» для r -го теста.

Каждый l -й метод дает независимую ненормированную оценку w_m^l m -му признаку ($l = \overline{1, L}$, $m = \overline{1, M}$) и, соответственно, оценку W_r^l r -му тесту, $r = \overline{1, N_T}$. Пусть оценки ВКП, полученные по l -му методу, находятся в интервале $\left[w_{\min}^l, w_{\max}^l \right]$. Поскольку в общем случае при $l \neq t$, $t = \overline{1, L}$, $w_{\min}^l \neq w_{\min}^t$, $w_{\max}^l \neq w_{\max}^t$, то при построении коллективной оценки возникает необходимость пересчета оценок ВКП в шкалах с различными пределами к единой шкале, имеющей наибольшее количество делений. Для этого оценку w_m^l в

шкале с пределами $[w_{\min}^l, w_{\max}^l]$ выразим через «эталонную» оценку w_m^* с пределами $[w_{\min}^*, w_{\max}^*]$: $\bar{w}_m^* = a_m^* w_m^* + b_m^*$, где коэффициенты пересчета и смещения a_m^* и b_m^* равны:

$$\begin{aligned} a_m^* &= (w_{\max}^l - w_{\min}^l) / (w_{\max}^* - w_{\min}^*), \\ b_m^* &= w_{\min}^l - w_{\min}^* (w_{\max}^l - w_{\min}^l) / (w_{\max}^* - w_{\min}^*). \end{aligned} \quad (4.18)$$

Таким образом, вес r -го теста по l -му методу, равен:

$$W_r^l = \sum_{m \in \{m_1, m_2, \dots, m_h, \dots, m_{g_r}\}} \bar{w}_m^l, \quad (4.19)$$

где $i_h \in \{1, \dots, M\}$, g_r – количество признаков в r -ом тесте.

Коллективную оценку W_r для r -го теста по всем методам на основе вычисленных оценок тестов по формуле (4.19), традиционно определяют на основе линейной свертки:

$$W_r = \sum_{l=1}^L c_r^l W_r^l, \quad (4.20)$$

где компоненты вектора коэффициентов $C_r = (c_r^1, c_r^2, \dots, c_r^L)$ выбирают пропорционально эффективности соответствующих методов, например, как вектор частот безошибочной классификации по рассматриваемым методам на контрольной выборке. В настоящей работе показаны (гл. 3) условия применимости линейной свертки (теоремы 3.1-3.4) и приведены примеры 3.1-3.3 несостоятельности ее применения. Поэтому вместо формулы (4.20) рекомендуется использовать формулы (4.21) (нелинейную свертку частных оценок, см. табл. 3.1):

$$W_r = \sum_{j=1}^L w_{rj}^l(r), \quad w_{rj}^l(r) = \sum_{l=1}^L c_l w_{rj}^l(r) = \sum_{l=1}^L c_l \frac{w_r^l}{w_r^l + w_j^l}, \quad w_r^l = \left(\prod_{l=1}^L a_{rl}^l \right)^{1/L}, \quad (4.21)$$

где величины $a_{km}^l = w_k^l / w_m^l$ – элементы МПС оценок по l -му методу.

Замечание. Привлечение экспертов для оценивания тестов широко используется для повышения достоверности принимаемых решений. Для сравнительной оценки двух подходов к определению весов тестов, вычислительного и экспертного, используемых в принятии решений в интеллектуальных системах, может быть использован функционал качества, аналогичный (4.17), минимум которого укажет на более эффективный подход при принятии итогового решения для рассматриваемой задачи распознавания.

Выводы. Трудность задачи определения весовых коэффициентов признаков обусловлена многими причинами [126, 132]: некорректность применения

процедур «усреднения» значений признаков в случае их взаимозависимости; возможная противоречивость (или пропуски) данных в случае их оценивания экспертами; большая размерность описания рассматриваемой задачи и прочие. С помощью вышеизложенных методов частично преодолены некоторые из перечисленных проблем.

Отметим, что весьма важным моментом в процессе выявления закономерностей и принятия решений в интеллектуальных системах является не только определение совокупности наиболее информативных (значимых) признаков (критериев), но и выявление отношений между ними. Приведенные методы позволяют количественно учесть вклад каждого признака в различающую способность теста в условиях независимости (методы 1 – 5) и возможной взаимозависимости (методы 6, 7) входящих в тест признаков.

Применение методов 6, 7 основанного на формализме мультимножеств, привело к возможности представления весового коэффициента рассматриваемого теста не в виде суммы ВКП признаков, что не корректно²⁵ в случае их взаимозависимости (пересечение соответствующих мультимножеств не пусто), а в виде суммы мощностей непересекающихся мультимножеств, составляющих тест.

В качестве показателя эффективности метода получения ВКП (в тестовом распознавании) и, соответственно, метода получения весовых коэффициентов тестов предложено использовать функционал качества, определяющий эмпирическую ошибку распознавания по каждому из рассматриваемых методов, что считается традиционным в теории распознавания образов. Выбирая коэффициенты пропорциональности соответственно эффективности метода на контрольной выборке, возможно построение коллективной оценки для весовых коэффициентов признаков (тестов).

Сформулирован ряд теорем. Теорема 4.1 устанавливает методологически важную связь между формулой (4.9), дающей значения весовых коэффициентов признаков на основе их разделяющей способности и оценкой главного собственного вектора, полученной на первом этапе метода 6. Тем самым показано, что метод 2 определения ВКП является частным случаем метода 6.

Теоремы 4.2 и 4.3 содержат обоснование подхода к формированию МПС признаков, позволяющего существенно сократить количество вычислений при подсчете весовых коэффициентов признаков с мерами относительной важности признака i над признаком j , определяемыми по формуле (4.11), учитывающими возможную взаимозависимость входящих в тест признаков.

С математической точки зрения все перечисленные методы применимы для вычисления ВКП и весовых коэффициентов тестов, но эффективность их

²⁵ Кини Р.Л., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения. – М.: Радио и связь, 1981. – 560 с.

использования существенно зависит от того, насколько оправданными являются предпосылки и предположения, лежащие в основе каждого из этих методов.

Приведенный иллюстративный пример дает основание полагать, что учет специфики задачи с целью выбора соответствующего (более подходящего) метода приводит к построению более точных оценок значений весовых коэффициентов признаков, используемых в интеллектуальных системах различного назначения.

Численное оценивание эффективности применения весовых коэффициентов проведено на реальных базах данных [62, 63] и показало результат снижения тестовой ошибки распознавания при использовании ВКП и ВКТ с 37% (без применения ВКП) до 0.07% (с применением ВКП и ВКТ на основе g-эталонных).

Результаты, представленные в 4-й главе, опубликованы в работах [21-23, 73, 77, 79, 88, 91-100, 171-173, 175].

«...нередко случается, что ловкой обработкой одного и того же материала можно выжать из него при помощи этого приема [методов группировки данных по интервалам] прямо противоположные заключения» /А.А. Чупров²⁶/

Глава 5. МОДЕЛЬ ВЫЯВЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ В ТЕОРЕТИКО-ИНФОРМАЦИОННОМ ПОДХОДЕ К РАСПОЗНАВАНИЮ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА

Информационный подход является математической основой многих методов [68, 130]. В частности, можно показать его согласованность с критерием релевантности, основанным на различии условных распределений признака в классах, и с ROC-анализом (Receiver Operator Characteristic), результатом которого является кривая, наиболее часто используемая для представления результатов бинарной классификации в машинном обучении. В работах [58, 140] энтропийный подход использован для оценки информативности признаков, используемых при распознавании образов (гл. 4), а также в задачах контроля и диагностики сложных систем [138], оценки рисков [155], а также при решении ряда других задач [45].

В данном разделе, во-первых, обоснована принципиальная возможность получения вектора оценок вероятностей состояний ДО и построена соответствующая итеративная процедура на основе максимального правдоподобия; во-вторых, построена процедура для принятия решения о номере состояния ДО на основе дивергенции Кульбака; в-третьих, поставлена задача поиска системы градаций квантования минимальной размерности, которая «наилучшим образом» выделяла бы информативность градаций в смысле их дискриминирующей сути и предложен соответствующий алгоритм.

Под информативностью понимается качественное свойство, характеризующее долю полезной информации в общем объеме СГК и выраженное количественно определенной мерой (энтропией, дивергенцией информации и др.).

5.1. Оптимальная оценка вероятности состояний сложного объекта на основе максимального правдоподобия

Введем обозначения: $p^i(t)$ – вероятность i -го состояния ДО Ω_i , $i = \overline{1, I}$, в момент времени t ; $p_v^i(t) = P_t(v / i)$ – вероятность попадания значения временного ряда (показателя, координаты ДО) в градацию v -го уровня при условии i -

²⁶ Ученый в области математики, статистики, экономики, педагог, 1874-1926.

го состояния ДО; $p_v(t) = \sum_{i=1}^I p^i(t) P_i(v/i)$ – вероятность градации v -го уровня в момент времени t . Вероятность $P_i(v/i)$ может быть оценена на обучающей выборке (или на основе экспертного анализа).

Далее рассмотрим случай, когда вероятности $p^i(t)$ и $P_i(v/i)$ остаются постоянными на интервале наблюдения ($n(t)$ -окно) $[t, t+n-1]$ (индекс t , где не искажается изложение, будем опускать).

Обозначим n_v , $\hat{p}_v = n_v/n$ – число появлений в последовательности $y(t, n) = \{y_t, y_{t+1}, \dots, y_{t+n-1}\}$ градации v -го уровня и частота появлений, соответственно, $\sum_{v \in U} n_v = n$.

Для оценки распределения $P_\Omega = (p^1, \dots, p^I)$ по опытным данным используем метод максимального правдоподобия (ММП), который приводится к одному из трех методов для трех распространенных законов распределения: методу наименьших квадратов (для с.в. ξ , распределенной по Гауссу), метод наименьших модулей (для с.в. ξ , распределенной по Лапласу), метод минимизации наибольшего уклонения (для с.в. ξ , распределенной равномерно).

Постановка задачи. Пусть дана обучающая выборка, по которой сформирован наиболее информативный домен U (множество уникальных значений градаций v) и система градаций квантования (СГК) $G = \{\kappa_v, v : v \in U\}$, $U = \{v_j : j = \overline{1, N_U}\}$ (G – мультимножество, см. определение 1.15), способы нахождения наиболее информативной системы градаций квантования будут ниже указаны. Требуется получить набор $P_\Omega = (p^1, \dots, p^I)$ вероятностей распознаваемых состояний ДО ($P_\Omega = P_\Omega(t)$), максимизирующий вероятность $p_y(P_\Omega)$ последовательности наблюдений $y(t, n) = \{y_{t-n}, y_{t-n+1}, \dots, y_t\}$.

Решение задачи. Рассмотрим простой вариант в предположении о статистической независимости появления градаций (признаков). Вероятность последовательности $y(t, n)$ с учетом введенных обозначений будет определяться по формуле:

$$p_y(P_\Omega) = \prod_{v \in U} \left(\sum_{i=1}^I p^i p_v^i \right)^{n_v}, \quad (5.1)$$

Параметры распределения $P_\Omega = (p^1, \dots, p^I)$ найдем из условия максимума величины $p_y(P_\Omega)$ или, что эквивалентно, минимума величины $-\log_2 p_y(P_\Omega)$:

$$\begin{aligned} -\log_2 p_y(P_\Omega) &= -\log_2 \prod_{v \in U} \left(\sum_{i=1}^I p^i p_v^i \right)^{n_v} = \\ &= -n \sum_{v \in U} \frac{n_v}{n} \log_2 \sum_{i=1}^I p^i p_v^i = -n \sum_{v \in U} \hat{p}_v \log_2 \sum_{i=1}^I p^i p_v^i. \end{aligned} \quad (5.2)$$

Обозначим

$$H(P_\Omega, P_G) = -\log_2 p_y(P_\Omega)/n = -\sum_{v \in U} \hat{p}_v \log_2 \sum_{i=1}^I p^i p_v^i, \quad (5.3)$$

Для дальнейшего изложения потребуется теорема из [48], приведенная ниже в немного измененной формулировке, и результаты из [164, 165].

Теорема 5.1. Если функция $f(\alpha) \in R$ – выпуклая вверх (вогнутая), где $\alpha = (\alpha_1, \dots, \alpha_k)$ – вектор вероятностей и частные производные $\partial f(\alpha)/\partial \alpha_k$ определены и непрерывны в области R , тогда необходимыми и достаточными условиями того, что вектор вероятностей α максимизирует $f(\alpha) \in R$ являются следующие выражения:

$$\partial f(\alpha)/\partial \alpha_k = \lambda, \forall k : \alpha_k > 0; \quad \partial f(\alpha)/\partial \alpha_k \leq \lambda, \forall k : \alpha_k = 0, \quad (5.4)$$

Сформулируем теорему, являющуюся теоретическим обоснованием условия корректности постановки задачи в форме (1.1) и содержащую уравнение для определения набора $P_\Omega = (p^1, \dots, p^I)$.

Теорема 5.2. При выполнении условия (1) $P_\Omega : \exists i = \overline{1, I} : p_i > 0$ (не все компоненты $P_\Omega = (p^1, \dots, p^I)$ одновременно равны нулю) существует единственный набор вероятностей $P_\Omega = (p^1, \dots, p^I)$, максимизирующий вероятность последовательности наблюдений $p_y(P_\Omega)$; при выполнении условия (2) $P_\Omega : \forall i = \overline{1, I} : p_i > 0$ единственный набор вероятностей $P_\Omega = (p^1, \dots, p^I)$ максимизирует вероятность последовательности наблюдений $p_y(P_\Omega)$ тогда и только тогда, когда для каждого $i = \overline{1, I}$ имеет место соотношение:

$$\sum_{v \in U} \hat{p}_v p_v^i \left(\sum_{i=1}^I p^i p_v^i \right)^{-1} = 1, \quad (5.5)$$

Доказательство теоремы 5.2. изложено в Приложении 5.

5.2. Итеративная процедура для выбора системы признаков на основе максимального правдоподобия

Для интервала наблюдения $[t, t+n-1]$, $y(t, n) = \{y_t, y_{t+1}, \dots, y_{t+n-1}\}$ итеративно строим ММП-оценку $\hat{P}_\Omega(t) = (\hat{p}^1, \dots, \hat{p}^I)$, значение которой участвует в принятии решения о метке (номере) состояния по алгоритму 5.1.

Выходом итеративной процедуры оцениваемого распределения $P_\Omega = (p^1, \dots, p^I)$ будем считать значения $\hat{P}_\Omega = (\hat{p}^1, \dots, \hat{p}^I)$, полученные с указанной точностью для СГК $G = \{\kappa_\nu, \nu : \nu \in U\}$, $U = \{\nu_j : j = \overline{1, N_U}\}$.

Согласно теореме 5.2 при выполнении 2-го условия $P_\Omega : \forall i = \overline{1, I} : p_i > 0$ имеют место уравнения (5.5), решения которых $P_\Omega = (p^1, \dots, p^I)$ максимизируют вероятность $p_G(P_\Omega)$, или, что эквивалентно, минимизируют (выпуклую) функцию $-\log_2 p_G(P_\Omega)/n = H(P_\Omega, P_G)$. Аналитически уравнение (5.5) не разрешимо, ниже излагается процедура, позволяющая с заданной точностью итеративно определить набор $\hat{P}_\Omega(t) = (\hat{p}^1, \dots, \hat{p}^I)$, максимизирующий вероятность $p_G(P_\Omega)$, а на ее основе предлагается осуществление выбора оптимальной СГК из нескольких назначенных экспертно или найденных на базе нижеизложенного алгоритма 5.2.

Алгоритм 5.1 определения набора $\hat{P}_\Omega(t) = (\hat{p}^1, \dots, \hat{p}^I)$ опирается на преобразование $P'_\Omega = F(P_\Omega)$:

$$p'^i = \sum_{\nu \in U} \frac{\hat{p}_\nu p^i P(\nu / i)}{\sum_{j=1}^I p^j P(\nu / j)}, \quad i = \overline{1, I}, \quad (5.6)$$

Лемма 5.1. Преобразование $P'_\Omega = F(P_\Omega)$ (5.6) обладает свойствами:

1) совокупность значений $\{p'^i, i = \overline{1, I}\}$, полученных по формуле (5.6), является вероятностным распределением;

2) $F : \Theta \rightarrow \Theta$, где Θ - замкнутое множество:

$$\Theta = \left\{ P'_\Omega = (p'^1, \dots, p'^I) \mid p'^1 \geq p^1, \dots, p'^I \geq p^I, \sum_{i=1}^I p'^i = 1 \right\};$$

3) при условии $P_\Omega : \forall i = \overline{1, I} : p_i > 0$ набор $P_\Omega = (p^1, \dots, p^I)$ из (5.6) единственный и является ММП-оценкой искомого распределения состояний ДО.

Алгоритм 5.1. Оценивание распределения состояний ДО на основе метода максимального правдоподобия

Вход: Наблюдения $y(t, n) = \{y_t, y_{t+1}, \dots, y_{t+n-1}\}$.

Выход: $\hat{P}_\Omega = (\hat{p}^1, \dots, \hat{p}^I)$ - вектор оценок вероятностей состояний $\Omega_i, i = \overline{1, I}$.

1: Формируем начальное приближение $\{p_0^i, i = \overline{1, I}\}$. Полагаем $l=1$.

2: К приближению $\{p_{l-1}^i, i = \overline{1, I}\}$ применяем преобразование (5.6), полу-

чаем
$$p_l^i = \frac{\hat{p}_v p_{l-1}^i P(v/i)}{\sum_{j=1}^I p_{l-1}^j P(v/j)}, \quad i = \overline{1, I}.$$

3: Вычисляем величину $h(P_l^i, P_{l-1}^i) = |H(P_l^i, P_G) - H(P_{l-1}^i, P_G)|$. Если значение $h(P_l^i, P_{l-1}^i) < \varepsilon$, то полагаем $\{\hat{p}^i, \hat{p}^i = p_l^i, i = \overline{1, I}\}$. В противном случае выполняем шаг 2.

Вернуть $\hat{P}_\Omega = (\hat{p}^1, \dots, \hat{p}^I)$.

Свойство алгоритма сформулировано в леммах 5.1, 5.2 и теореме 5.3.

Лемма 5.2. Функция $H(P_l^i, P_G)$ не возрастает по l , или $h(P_l^i, P_{l-1}^i) = H(P_l^i, P_G) - H(P_{l-1}^i, P_G) \leq 0$.

Теорема 5.3. Последовательность $\{\hat{P}_l, l \geq 1\}$, $\hat{P}_l = (\hat{p}_l^1, \dots, \hat{p}_l^I)$, определяемая по алгоритму 5.1 и формуле (5.6), сходится к ММП-оценке (5.5).

1.1. Теоретико-информационный подход к выявлению закономерностей в задаче распознавания состояний динамического объекта

Будем использовать постановку задачи (1.1) распознавания состояний динамического объекта (гл. 1).

Процедура распознавания состояний ДО состоит из 2-х этапов: итогом первого этапа является совокупность информативных (характеристических) признаков в виде оценок условных распределений значений уровней квантования в разных состояниях ДО. Итогом второго этапа является принятие решения о принадлежности наблюдаемого сигнала одному из состояний ДО.

1 этап. Определение характеристических признаков.

1.1. Формируем СГК с наибольшей информативностью на обучающей выборке (см. ниже алгоритм 5.2).

1.2. Определяем совокупность информативных (характеристических) при-

знаков, являющимися оценками распределения вероятностей $p^i = \{p_\nu^i, \nu = \overline{0, N_U}\}$, $i = \overline{1, I}$ логарифмов уровней ВР, характерных для каждого состояния ДО по формуле (p_ν^i - оценка вероятности ν -й градации для i -го состояний ДО (образа), $i = \overline{1, I}$) $p_\nu^i = N_\nu^i / N^i$ (N^i - общее число дискретных отсчетов для i -го состояний ДО, N_ν^i - число встреч сигнала ν -го уровня для i -го состояний ДО). Усредняем значение вероятностей по числу доступных реализаций.

2 этап. Распознавание состояний ДО (принятие решения).

2.1. По результатам наблюдений $y(t, k) = (y(t-k), y(t-k+1), \dots, y(t))$ (k -окно) вычисляем соответствующие относительные частоты $p_\nu = N_\nu / n$, $\nu \in U$ (N_ν - число появлений признака (градации) ν).

2.2. Вычисляем информационное расстояние между «эталонными» оценками p^i условных вероятностных распределений с.в. - числа встреч определенной градации сигнала в каждом состоянии ДО и полученными в результате наблюдений $p = \{p_\nu, \nu \in U\}$ по формулам:

$$D(p^i, p) = d\left(p^i, \frac{p^i + p}{2}\right) + d\left(p, \frac{p^i + p}{2}\right), \quad (5.7)$$

$$d\left(p^i, \frac{p^i + p}{2}\right) = c \sum_{\nu \in U} p_\nu^i \log_2 \frac{2p_\nu^i}{p_\nu^i + p_\nu},$$

$$d\left(p, \frac{p^i + p}{2}\right) = c \sum_{\nu \in U} p_\nu \log_2 \frac{2p_\nu}{p_\nu^i + p_\nu},$$

где $p^{i*} = (p^i + p) / 2$ - распределение вероятностей, которое получается усреднением для каждой пары соответствующих вероятностей из исходных p^* и p^i ; c - нормирующий коэффициент, обеспечивающий выполнение неравенства $0 \leq d(p^i, p) \leq 1$; сумма $d(p^i, p^{i*}) = \sum_{\nu \in U} p_\nu^i \log_2 \frac{p_\nu^i}{p_\nu}$ - оценка расстояния Кульбака - Лейблера (меры, характеризующей взаимную энтропию 2-х вероятностных распределений). Величина $d(p^i, p)$ неотрицательна и равна 0, если равны сравниваемые распределения, и дает в битах взаимную информацию.

2.3. Принимаем решение в пользу того состояния ДО, для которого величина $D(p^i, p)$ является наименьшей: $i^* = \arg \min_{i=\overline{1, I}} (D(p^i, p))$.

Замечание. Мера различий распределений (5.7) введена как альтернатива расстоянию Кульбака - Лейблера, которое не является метрикой (несимметрич-

но и не удовлетворяет аксиоме треугольника).

Утверждение 5.1. Функция $D(p^i, p)$, определяемая по формуле (5.7), удовлетворяет всем свойствам расстояния.

Доказательство проводится непосредственной проверкой свойств расстояния (метрики).

В табл. 5.1 приведены результаты моделирования, где указана зависимость эффективности распознавания (%), времени распознавания для 10000 объектов (10^{-5} сек) и времени обучения (10^{-6} сек) от параметров шума. В табл. 5.1 используются обозначения: $U[a; b]$ – равномерный шум на интервале $[a; b]$; $N(m; \sigma^2)$ – нормальный шум с параметрами $(m; \sigma^2)$; ОВ - обучающая выборка; КВ - контрольная выборка; б/ш - без шума. Число квантов в Ω_1 – число градаций в минимальном по уровню состоянию ДО, в примере это Ω_1 .

Таблица 5.1. Распознавание состояний ДО на основе информационной метрики

Число квантов в Ω_1 / Величина шага квантования		Эффективность распознавания, %				Время распознавания	Время обучения
		ОВ б/ш	КВ1 $U[-320;320]$	КВ2 $U[-640;640]$	КВ3 $N(0,100)$		
5	5.746	0.6851	0.6835	0.6853	0.6801	406	28
10	2.873	0.8606	0.8536	0.8685	0.7774	690	28
40	0.718	0.9616	0.9647	0.9661	0.8456	701	50
100	0.287	0.9398	0.9387	0.9362	0.8917	1003	65

Таким образом, информационный подход с применением идеи распознавания состояний как образов и применения метрики, основанной на дивергенции Кульбака, дает возможность распознавания состояний динамического объекта за приемлемое (реальное) время с приемлемым качеством.

Заметим, однако, существенное различие в качестве распознавания состояний ДО в зависимости от величины шага квантования (2-й столбец в табл. 5.1). Как известно, метод задания гистограмм является тонкой работой, требующей чаще всего ручной настройки, и поэтому автоматическое и адаптивное задание гистограмм (здесь системы градаций квантования (СГК)) для анализируемого процесса не является практически решенной задачей.

5.4. Оценка информативности квантования временного ряда в задаче распознавания состояний динамического объекта

На подготовительном этапе распознавания при формировании характеристических признаков состояний ДО естественно поставить задачу нахождения

такого разбиения системы исходных измерений на новые градации (подмножеств значений из интервала $[0, \kappa_U]$) (κ_U - максимальный уровень квантования), которое наилучшим образом выделяет дискриминирующую суть градаций как признаков распознавания состояний ДО. Внешними критериями оценки системы признаков являются эффективность распознавания и время распознавания (гл. 2), внутренними критериями оценки модели квантования являются энтропия и дивергенция информации Кульбака.

Обзор существующих методов [125, 151, 159] показывает, что к настоящему времени в известной литературе не приводится эффективных (для определения характеристик ДО) методов обоснования частоты дискретизации, её стохастических характеристик, от которых непосредственно зависят связанные с ней интерполяция и экстраполяция. Так, известной теореме отсчётов Котельникова (Найквиста, Шеннона) присущи ряд свойств, которые не позволяют считать её вполне применимой для сложных зашумленных ДО. Перечислим некоторые из них: неограниченность спектров реальных сигналов, невозможность создания идеального фильтра, сложность расчётов при восстановлении функции членами ряда Котельникова, неопределённость интервала между отсчётами в силу произвольности граничной частоты, невозможность определения стохастических характеристик погрешностей при дискретизации.

5.4.1. Практическая реализация адаптивного выбора эффективной системы признаков

Определим систему признаков для распознавания состояний ДО на основе формализма мультимножеств.

Определение 5.1. Стохастическое мультимножество (определения 1.15) $G = \{k_1 \bullet v_1, k_2 \bullet v_2, \dots, k_{N_U} \bullet v_{N_U}\}$, задаваемое основанием (доменом) $U = \{v_j : j = \overline{1, N_U}\}$ – кортежем уникальных элементов (градаций) и кортежем кратностей $[G] = (k_1, k_2, \dots, k_{N_U})$ элементов основания (первичная спецификация G), назовем стандартной (равномерной, линейной) системой градаций квантования (СГК) G , если градации квантования (уровни ВР) получены по линейному закону: $v_j = s_j \cdot h$, где s_j – целое число, h - шаг квантования, равный частному от деления диапазона значений ВР на заданное число отрезков.

Суть адаптивного подхода к решению проблемы оптимального квантования заключается в выборе СГК минимальной размерности (мощности домена U' , $|U'| = N_{U'}$, или размерности мультимножества G'), которая сохраняла бы информативность исходной «частотной картины» состояния и обладала бы желательными свойствами (меньшее время распознавания с приемлемым качеством).

Следующее определение опирается на новое понятие стохастического мета-мультимножества.

Определение 5.2. Стохастическое мультимножество $G' = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{N'}\}$, задаваемое элементами-мультимножествами $\mathcal{G}_j : j = \overline{1, N'}$ (мета-градациями) $\mathcal{G}_j = \{k_{j1} \bullet v_{j1}, k_{j2} \bullet v_{j2}, \dots, k_{jn_j} \bullet v_{jn_j}\}$, состоящими из элементов - уровней градаций $v_j \in U : j \in \{1, \dots, N_U\}$, назовем адаптивной (неравномерной, нелинейной) системой градаций квантования (СГК, мета-мультимножеством) G' . Мультимножество \mathcal{G}_j является объединением элементов равномерного квантования $\{v_l, v_l \in [c_j, c_{j+1}]\}$, где c_j, c_{j+1} – границы «карманов», определяемые в соответствии с максимальным значением информационной меры.

Другими словами, СГК G' получается разбиением диапазона значений СГК G на некоторое заранее не заданное количество интервалов, границы которых заранее не определены.

На рис. 5.1 показана нежелательная особенность равномерного квантования, приводящая к существенному падению качества распознавания состояний ДО. График точности распознавания одного из 4-х состояний сигнала напряжения асинхронного двигателя в зависимости от разного числа градаций показывает резкое падение точности распознавания. Анализ показал, что в одном из состояний на временном ряде на отдельную градацию приходится $\sim 10\%$ значений от всего объема значений ВР, приходящихся на это состояние (ввиду логарифмирования исходного ряда). В связи с этим число ошибок распознавания для объектов из данного состояния сразу увеличивается на ~ 3000 (33%, т.к. всего генерировалось в данном эксперименте 10000 объектов). Для отдельных реализаций сигнала (как СП) число ошибок при распознавании одного состояния составляли 100% от общего числа ошибок.

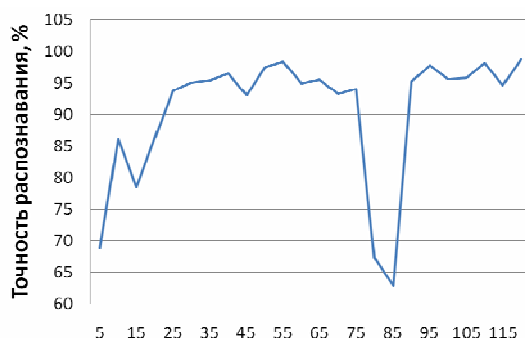


Рис. 5.1. Результат неудачного выбора шага равномерного квантования, приведший к падению качества распознавания состояний ДО

Поскольку высокий уровень шума вносит существенную погрешность в вычисления при равномерном квантовании, то в этих условиях формирование (диагностических) признаков для принятия решения о состоянии ДО может проводиться только для временного ряда с низким уровнем зашумленности.

Постановка задачи. Пусть дана исходная априорная классификация состояний ДО $P_\Omega = (p^1, \dots, p^I)$ на основе равномерного квантования СГК $G_0 = \{\kappa_\nu, \nu : \nu \in U\}$, $U = \{\nu_j : j = \overline{1, N_U}\}$, построенной на обучающей выборке $y'(t_0, N)$ анализируемого показателя (характеристики).

Требуется сформировать такую систему признаков - $G_1 = \{\mathcal{G} : \mathcal{G} = \overline{1, N_{1U}}\}$ (такое разбиение исходной СГК $G_0 = \{\nu : \nu \in U\}$ на новые мета-градации), которая приводила бы к максимуму информационной меры относительно набора вероятностей распознаваемых состояний ДО $P_\Omega = (p^1, \dots, p^I)$.

В качестве меры информативности системы признаков G_1 для достижения цели - распознавания состояний ДО – выберем:

1) минимум энтропии вероятностного распределения состояний ДО

$$H(P_\Omega) = -\sum_{i=1}^I p_i \log p_i ;$$

2) максимум информационной меры на основе дивергенции С. Кульбака.

Решение задачи. Способ выбора СГК на основе энтропийного критерия. Понятно, что чем больше бит информации принесет каждая градация (как признак), тем эффективнее СГК. Энтропия случайной величины ν (показателя), имеющей конечное число значений $\{\nu : \nu = \overline{\nu_1, \nu_{N_U}}\}$ с распределением вероятностей $\{p_j\}$ определяется как математическое ожидание собственной информации $I(\nu)$: $H(\nu) = M(I(\nu)) = -\sum_{j \in U} q_j \log q_j$ (U - множество значений с.в. ν , или домен

в терминологии формализма множеств). Чем больше значение энтропии, тем более широкую область занимает частотная гистограмма, и значения компонент варьируются в более широком диапазоне, а это значит, что оно содержит больше информации. Поэтому одним из критериев целенаправленного преобразования СГК (или выбора из нескольких) является максимизация энтропии распределения градаций как признаков (q_j – вероятность j -й градации в СГК). Известно, что из всех статистических функций распределения, определенных на конечном интервале, максимальной энтропией обладает равномерное распределение. Следовательно, целью преобразований будет «выравнивание» имеющейся выборки исходных градаций, полученных в результате стандартного кванто-

вания, за счет «подходящих» преобразований (объединений) исходных градаций в мета-градации.

Число возможных разбиений исходного множества уровней квантования на k непустых групп определяется величиной $\sum_{j=0}^k (-1)^j C_k^j (k-j)^{N_U}$, где C_k^j - число сочетаний из k по j . В связи с этим задача нахождения оптимального разбиения является NP -сложной, решение которой весьма затруднительно с вычислительной точки зрения даже при современных технологиях [19, 42, 49]. Этот факт говорит о важности создания эффективных алгоритмов поиска оптимальной СГК.

Изложим кратко простой эвристический алгоритм создания эффективной (в смысле заданных критериев качества) системы уровней квантования на основе сведения исходного равномерного квантования к адаптивному (неравномерному).

Диапазон (размах) допустимых значений показателя разбивается на N' отрезков (по числу «карманов»), длины которых пропорциональны числу соответствующих градаций в исходной СГК (выборке), попадающих в «карман»: $\gamma_j = |\mathcal{G}^j| = N_U^j / N_U$ (N_U - общее число градаций в СГК G , N_U^j - число градаций, попадающих в мета-градацию (множество) \mathcal{G}^j , $|x|$ - мощность множества x).

Общая теоретическая идея упрощенного алгоритма для формирования адаптивной СГК (см. алгоритм 5.2), применение которого проверено на практике в реальном времени (~мкс), следующая.

По обучающей выборке в виде ВР (дискретной реализации наблюдаемой характеристики ДО) формируются множества U (множество уникальных градаций) и $[G]$ (кортеж кратностей градаций как спецификация стохастического мультимножества G). Если установлено, что качество равномерного квантования неудовлетворительно (что свидетельствует о неравномерном исходном распределении значений показателя), то выделение «карманов» осуществляется, исходя из условия равенства площадей фигур, образованных вертикальными сечениями от граничных значений «карманов» до кривой, огибающей гистограмму оценки функции плотности распределения $f(x)$.

Практически эта идея воплощается при формировании новой СГК (метамультимножестве G'), которая на основе множеств U и $[G]$ формируется по принципу: образуется новый уровень квантования, если выполняется хотя бы одно из двух условий: а) расстояние между текущим и предыдущим уровнями квантования превосходит пороговое значение Δ_{\max} ; б) разность между «встречаемостью» соседних (отсортированных) значений градаций превосходит b_{\max} (группирование пустых или мало заполненных интервалов). Величины Δ_{\max} и b_{\max} регулируются параметрами квантования θ и h_0 .

При таком способе преобразования исходных градаций все выделенные признаки-множества \mathcal{G}_j будут нести примерно одинаковую информационную нагрузку, т.е. перевод признаков-градаций, полученных в равномерной шкале, в систему признаков - мета-градаций, будет сопровождаться наименьшими потерями информации.

Алгоритм¹ 5.2 Адаптивный выбор уровней квантования

Вход: временной ряд $x(t)$, $t \in [t_0, T]$; параметры регуляризации квантования θ, h_0 .

Выход: SGK $G' = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{N'}\}$, $\mathcal{G}_j = \{k_{j1} \bullet v_{j1}, k_{j2} \bullet v_{j2}, \dots, k_{jn_j} \bullet v_{jn_j}\}$,

$v_j \in u_j : j \in \{1, \dots, N_U\}$, $u_j = \{v_l, v_l \in [c_j, c_{j+1}]\}$, $[G'] = (k_1, k_2, \dots, k_{N'})$.

1. Вычислить максимальное и минимальное значения анализируемого временного ряда x_{\max} и x_{\min} .

2. Сформировать исходное множество уникальных градаций $U = (x_1, x_2, \dots, x_n)$.

3. Подсчитать частоты для каждого элемента множества U . Результаты поместить в массив $\{n_x, x \in U\}$, упорядоченный по возрастанию значений элементов множества U .

4. Вычислить максимальный шаг квантования: $\Delta_{\max} = h_0(x_{\max} - x_{\min})/\theta$; параметр различения уровней квантования («перепадов» встречаемости): $b_{\max} = \max(|U|/\theta, 1)$.

5. Сформировать множество новых уровней квантования: положить $U' = \{x_{\min}\}$, нижнюю границу текущего «кармана» (кванта), равной $s_0 = x_{\min}$.

6. Положить $n_x(j)$ - частота текущего кванта j , $n_x(0) = 0$;

7. Выполнять цикл: $i = 1$ пока $i < n_x$.

a. Если $|n_x(j) - n_x(j-1)| > b_{\max}$ или $|x_j - s_0| > \Delta_{\max}$, то добавить во множество U' новый элемент $u_1 = 0.5(x_{j-1} + x_j)$ и обновить $s_0 = u_1$.

b. Увеличить $j = j + 1$. Если $j < n_x$, то переход на шаг 7.a, иначе переход на шаг 8.

8. Добавить во множество U' новый элемент x_{\max} : $U' := U' + x_{\max}$.

Вернуть: сформированное множество (набор уровней квантования U').

Изложим способ выбора SGK, гарантирующий «наилучшую» дискриминирующую суть новой системы признаков.

Способ выбора SGK на основе максимизации дивергенции Кульбака. Обо-

¹ Автор алгоритм Ю.Р. Цой, доцент ТПУ, осуществивший вычисления для главы 5.

значим через p_{dv}, q_{dv} - условные вероятности попадания значения градации $v \in U$ в d -й новый диапазон («карман») ($d = 1, 2, \dots, k, k > k_0, k_0$ задано) для состояний Ω_1 и Ω_2 , соответственно. Тогда для двух заданных состояний в качестве наилучшего разбиения исходного диапазона значений на k отрезков выбирается такое, которое максимизирует значение дивергенции Кульбака:

$$G_1 = \arg \min_G D(\Omega_1 : \Omega_2; G) = \arg \min_{k, U} \sum_{d=1}^k \sum_{v \in U} (p_{dv} - q_{dv}) \log(p_{dv}/q_{dv}).$$

В случае $I > 2$ состояний максимизируется величина:

$$D(G) = \sum_{i=1}^I \sum_{j=1}^I D(\Omega_i : \Omega_j; G). \text{ Вероятности, дивергенция информации Кульбака,}$$

дисперсия дивергенции оцениваются по данным обучающей выборки по формулам:

$$\begin{aligned} \hat{p}_v^l &= n_v^l / N_i^l, \quad \hat{q}_v^l = m_v^l / N_j^l, \\ \hat{D}_l &= \sum_{v \in U_l} (\hat{p}_v^l - \hat{q}_v^l) \log \frac{\hat{p}_v^l}{\hat{q}_v^l} = \sum_{v \in U_l} \left(\frac{n_v^l}{N_i^l} - \frac{m_v^l}{N_j^l} \right) \cdot \log \frac{n_v^l N_j^l}{m_v^l N_i^l}, \\ S(\hat{D}_l) &= \frac{1}{N_1^l} \left[\sum_{v \in U_l} \hat{p}_v^l \left(\log \frac{\hat{p}_v^l}{\hat{q}_v^l} + \frac{\hat{p}_v^l - \hat{q}_v^l}{\hat{p}_v^l} \right)^2 - \left(\sum_{v \in U_l} \hat{p}_v^l \log \frac{\hat{p}_v^l}{\hat{q}_v^l} \right)^2 \right] + \\ &+ \frac{1}{N_2^l} \left[\sum_{v \in U_l} \hat{p}_v^l \left(\log \frac{\hat{p}_v^l}{\hat{q}_v^l} + \frac{\hat{p}_v^l - q_v^l}{\hat{q}_v^l} \right)^2 - \left(\sum_{v \in U_l} q_v^l \log \frac{\hat{p}_v^l}{\hat{q}_v^l} \right)^2 \right], \end{aligned} \quad (5.8)$$

соответственно (верхние скобки, если это не искажает изложения, будем опускать), где N_i^l, N_j^l - объемы обучающей выборки для СДС Ω_i и Ω_j , соответственно, $i, j = \overline{1, I}, U_l$ - число уникальных градаций в G_l .

В [112] показано, что для $N_i^l \gg 1, N_j^l \gg 1$ оценка дивергенции информации \hat{D}_l является асимптотически несмещённой и сходится к D_l в средне квадратичном смысле.

На основании \hat{D}_l и $S(\hat{D}_l)$ возможно решение следующих задач:

-выбора СГК с наибольшей различающей способностью (информативностью) по максимуму дивергенции информации;

-получения количественной оценки различающей способности двух СГК по отношению их дивергенций информации, то есть по $\hat{D}_{l_1} / \hat{D}_{l_2}$.

-проверки гипотезы $D(G_1) = D(G_2)$ (две СГК G_1 и G_2 обладают одинаковой информативностью) при альтернативе $D(G_1) \neq D(G_2)$ на основе статистики

$$T = \frac{D(G_1) - D(G_2)}{\sqrt{\langle S(D(G_1)) \rangle + \langle S(D(G_2)) \rangle}}; \text{ принимается альтернативная гипотеза}$$

$D(G_1) \neq D(G_2)$, если выполнится условие $|T| \geq g_\alpha$, g_α – пороговое значение по уровню значимости α для стандартного нормального распределения;

–определения доверительного интервала для D_l по доверительному уровню $1-\alpha$ по формуле $D_l = \hat{D}_l \pm g_\alpha \sqrt{S(\hat{D}_l)}$, $l=1, 2$;

–вынесения решения о надёжности выбора СГК (см. ниже утверждение) в смысле оценки объёма обучающей выборки, гарантирующей требуемую верхнюю границу дисперсии $S(\hat{D}_l) < s_0$.

Обозначим в формуле (5.14) через κ_1 и κ_2 следующие величины:

$$\kappa_1 = \sum_{v \in U_l} p_v^l \left(\log \frac{\hat{p}_v^l}{\hat{q}_v^l} + \frac{\hat{p}_v^l - q_v^l}{\hat{q}_v^l} \right)^2 - \left(\sum_{v \in U_l} \hat{p}_v^l \log \frac{\hat{p}_v^l}{\hat{q}_v^l} \right)^2,$$

$$\kappa_2 = \sum_{v \in U_l} p_v^l \left(\log \frac{\hat{p}_v^l}{\hat{q}_v^l} + \frac{\hat{p}_v^l - q_v^l}{\hat{q}_v^l} \right)^2 - \left(\sum_{v \in U_l} q_v^l \log \frac{\hat{p}_v^l}{\hat{q}_v^l} \right)^2.$$

Утверждение 5.2. Для вынесения решения о надёжности выбора СГК в смысле выполнения неравенства $S(\hat{D}_l) < s_0$ необходимо выполнение условия на объём обучающей выборки в системе признаков G_i : $N_1^l > \kappa_1/s_0$, $N_2^l > \kappa_2/s_0$.

Доказательство утверждения следует из представления выражения для $S(\hat{D}_l)$ в виде: $S(\hat{D}_l) = \frac{\kappa_1}{N_1^l} + \frac{\kappa_2}{N_2^l}$. Из требуемого ограничения $S(\hat{D}_l) < s_0$ следует

условие $\frac{\kappa_1}{N_1^l} + \frac{\kappa_2}{N_2^l} < s_0$. Разрешив уравнение $\frac{\kappa_1}{N_1^l} + \frac{\kappa_2}{N_2^l} = s_0$ относительно N_2^l и

осуществив несложные преобразования:

$$N_2^l = \frac{\kappa_2 N_1^l}{s_0 N_1^l - \kappa_1} = \left(\frac{\kappa_2 N_1^l}{s_0 N_1^l - \kappa_1} \right) \frac{s_0}{s_0} = \frac{\kappa_2}{s_0} \left(\frac{N_1^l s_0}{N_1^l s_0 - \kappa_1} \right) = \frac{\kappa_2}{s_0} \left[1 + \frac{\kappa_1}{N_1^l s_0 - \kappa_1} \right],$$

получим соотношение: $N_2^l = \frac{\kappa_2}{s_0} + \frac{\kappa_2}{s_0} \cdot \frac{\kappa_1}{N_1^l \cdot s_0 - \kappa_1}$ – уравнение гиперболы относительно переменной N_1^l . Геометрически множество точек, удовлетворяющее требуемому неравенству $\frac{\kappa_1}{N_1^l} + \frac{\kappa_2}{N_2^l} < s_0$, находится в области $N_1^l > \kappa_1/s_0, N_2^l > \kappa_2/s_0$. Утверждение доказано.

В силу случайного характера обучающей выборки возможно получение нескольких «похожих» в смысле информативности СГК в ходе анализа временного ряда. Выбор «наилучшей» можно производить решением многокритериальной задачи на основе имеющихся данных по нескольким частным критериям

на основе ММАИ (алгоритм 2.3, гл.3), например, качество распознавания (f_1) состояний, время распознавания (f_2), полученных на контрольной выборке, размерность мультимножества G' (f_3) и другие, продиктованные спецификой решаемой задачи. Оценка по 1-му критерию количественно выражается в числе правильно распознанных объектов (значений ВР) относительно общего числа предъявленных к распознаванию. Отличительной особенностью применяемого для этой цели в данной работе корректной модификации известного метода парных сравнений является возможность анализа динамически изменяющихся сравниваемых наборов альтернатив (здесь, СГК) и гарантированное нахождение оптимума (наилучшее решение будет доставлять максимум используемой в ММАИ свертки критериев – функции относительного сходства).

5.5. Численное сравнение эффективности равномерного и адаптивного квантования при распознавании состояний ЭМС

Адаптивное квантование значений ВР, приводящее к значительному выигрышу во времени работы алгоритма распознавания без существенного изменения в качестве (рис. 5.2), является, по сути, способом, способствующим наиболее точному определению закономерностей (состояний) в анализируемом ВР. Так, из рис. 5.2 следует, что при числе признаков (градаций квантования), начиная с 40 (в минимальном по разнообразию уровней квантования состоянии) точность распознавания при адаптивном квантовании (АК) и стандартном равномерном квантовании (СК) выравнивается, а длительность распознавания при СК более, чем при АК, в 20-25 раз.

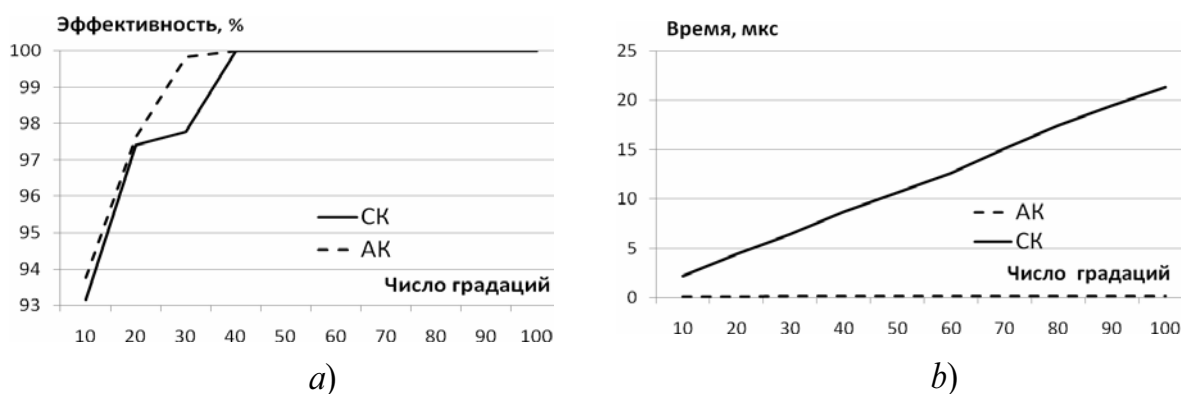


Рис. 5.2. Точность распознавания в зависимости от значения уровня квантования - а); б) - время распознавания (мкс, для 10^4 объектов, без учета времени на обучение) при адаптивном (АК) и стандартном квантовании(СК)

Покажем теперь, что процедура предварительной фильтрации для распознавания состояний ДО не является критичным звеном в распознавании сигналов (в смысле времени и эффективности распознавания (доли правильно распознанных объектов)). Исследование проводилось в условиях разных метрик (Евклида (ЕМ), Хемминга (ХМ), информационной обобщенной метрики (ИнфОМ)) (формула (5.7)) и при различных типах и уровнях шумов.

В табл. 5.2, 5.3 даны показатели эффективности и времени распознавания при стандартном (СК) и адаптивном выборах шага квантования (АК) и использовании процедуры распознавания на основе трех указанных метрик. В 1-й строке (для каждого типа сигнала) указан худший результат (в процентах), во 2-й строке – лучший, параметр квантования $h = 0.05$, число уровней квантования указано для минимального значения амплитуды сигнала, время распознавания указано для 1-го объекта (измерения) в мкс без учета времени на обучение и фильтрацию.

Таблица 5.2. Эффективность распознавания при стандартном и адаптивном выборах шага квантования по трем метрикам

Тип сигнала для распознавания	Эффективность распознавания (СК), %			Эффективность распознавания (АК), %		
	ИнфОМ	ЕМ	ХМ	ИнфОМ	ЕМ	ХМ
Без шума	86.48	84.73	86.71	78.81	78.00	65.38
	99.68	97.78	99.71	99.49	99.46	96.92
С шумом, 30%	76.19	76.05	76.4	73.62	80.08	77.05
	95.49	96.52	95.27	81.91	81.91	80.86
С фильтрацией	82.67	71.81	81.47	84.98	82.93	79.83
	94.49	95.13	94.24	81.6	82.93	86.49

Таблица 5.3. Время распознавания при стандартном и адаптивном квантовании по трем метрикам, $\bar{\Phi}$ - без фильтрации; Φ - с фильтрацией

Диапазон числа шагов квантования	Диапазон времени распознавания (СК)			Диапазон времени распознавания (АК)		
	ИнфОМ	ЕМ	ХМ	ИнфОМ	ЕМ	ХМ
10-50 $\bar{\Phi}$ Φ	0.167	0.050	0.026	0.026	0.013	0.008
	0.795	0.199	0.109	0.046	0.016	0.008
	1.267	0.510	0.321	0.225	0.123	0.090
	6.793	1.999	1.121	0.541	0.114	0.091
60-100 $\bar{\Phi}$ Φ	0.951	0.951	0.124	0.026	0.013	0.008
	1.583	1.583	0.209	0.047	0.017	0.009
	8.951	9.957	1.144	0.423	0.216	0.098
	14.583	15.581	2.249	0.749	0.245	0.099

Выводы по результатам моделирования на ЭВМ. Сравнение точности и времени распознавания с предварительной фильтрацией исходных данных и без фильтрации (на данных модели ЭМС) показало:

- 1) время распознавания с фильтрацией больше на один-два порядка;
- 2) точность распознавания при использовании стандартной процедуры квантования для данных без фильтрации выше, чем точность распознавания для данных с фильтрацией, примерно на 1%;
- 3) точность распознавания для адаптивной процедуры квантования при использовании фильтрации в среднем выше на 3-5%, чем без фильтрации. При параметре квантования, равном $h = 0.01$, точность распознавания отфильтрованного сигнала на 10% выше, чем сигнала с шумом;
- 4) для сигнала с шумом и отфильтрованного сигнала чуть более точно (примерно на 1%) работает распознавание с использованием метрики Евклида, по сравнению с использованием ИнфОМ и Хемминга метрик. При адаптивной процедуре квантования для отфильтрованного сигнала с метриками ИнфОМ и Хемминга точность распознавания выше на 2-4%, чем для метрики Евклида.

Среди различных алгоритмов распознавания на основе эталонов, подробно изложенных и экспериментально исследованных в [80, 84, 86] для решения обсуждаемой в этой работе задачи, лучше всего показывают себя алгоритм с использованием FriS-функции и алгоритм по методу ближайших соседей. При этом алгоритм с использованием FriS-функции лучше работает для сигнала без шума, а алгоритм по методу ближайших соседей лучше работает для сигнала с шумом и отфильтрованного сигнала.

Исследование по разным уровням и типам распределений шумов показало, что для равномерного (10%-го шума) увеличение числа градаций способствует повышению точности распознавания до некоторого порога (небольшого, порядка 30-ти градаций), начиная с которого точность распознавания существенно падает (на ~20%). Возможно, это связано с переобучением алгоритма распознавания (при увеличении числа градаций алгоритм «обучается на шуме»), т.к. замечено, что при малом уровне шума этот эффект отсутствует.

Исследование по разным 6-ти видам моделей (нелинейных и нестационарных) показало, что адаптивное квантование, сохраняя качество распознавания, приводит к существенно меньшему времени распознавания по сравнению со стандартным квантованием (табл. 5.3). Процедура фильтрации в тех ситуациях, где распознавание состояния ДО в реальном режиме времени является самостоятельной задачей, не несет основополагающей нагрузки, хотя на ряде моделей на фоне уменьшения шума алгоритмы распознавания показывают лучшие результаты (1%-5%).

Выводы. В главе 5 рассмотрен теоретико-информационный подход к распознаванию состояний ДО по наблюдаемым характеристикам с целью возможного использования этого знания для осуществления мониторинга ДО, основан-

ный на одновременном использовании методов распознавания образов и теории информации. Возможность распознавания состояний ДО за приемлемое время продемонстрирована на примере распознавания состояний ЭМС.

Экспериментально показано, что в тех задачах, где распознавание состояния сложного динамического объекта в реальном режиме времени является самостоятельной целью, процедура фильтрации не является критичным звеном в алгоритме распознавания. Предложенные алгоритмы могут быть использованы при организации робастного и адаптивного управления сложными ЭМО.

Результаты, представленные в 5-й главе, опубликованы в работах [80-87, 89, 90, 144-146].

«Правильнее будет сказать, что для данной точности (сколь угодно большой, но конечной) можно всегда указать такой промежуток времени, для которого становится невозможным сделать предсказания. И этот промежуток (и в этом вся соль) не так уж велик» /Р. Фейнман²⁸/

Глава 6. ОЦЕНКА ЭФФЕКТИВНОСТИ АЛГОРИТМОВ РАСПОЗНАВАНИЯ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА

Выше было отмечено, что статистические оценки качества [42] обучения для алгоритмов с бесконечной емкостью (п. 1.2) становятся тривиальными. Комбинаторные функционалы качества [47] обосновывают общий подход к выбору лучшего метода обучения из фиксированного набора по критерию скользящего контроля на заданной конечной выборке прецедентов, но, как и в [42], требуют решения дополнительной непростой задачи – оценки емкости семейства алгоритмов. Ниже предлагается вероятностный подход к оценке надежности распознавания состояний стохастической системы (эффективности алгоритма), под которой будем понимать способность алгоритма распознавания обеспечить заданный уровень доли правильно распознанных объектов из контрольной выборки $\Phi = q / m$, где q – число правильно распознанных объектов из выборки, m – общее число объектов в выборке.

Основу данного подхода к оценке качества распознавания составляет исходная информация о частотах встречаемости признаков (оценки вероятностей проявления признака) для каждого из классов (образов). Для ряда проблемных областей, например, в медицине, геологии, биологии, для некоторых классов (заболеваний, месторождений и др.) на основе статистического анализа известны вероятности проявлений совокупности признаков (симптомов, синдромов).

Предполагается, что объекты из O описываются некоторой системой признаков $Z = \{z_1, \dots, z_n\}$. Имеется набор O_1, \dots, O_m обучающих объектов (прецедентов) из $O = \{O_1, \dots, O_m\}$, о которых известно, каким классам они принадлежат. Требуется по предъявленному набору значений признаков (b_1, \dots, b_n) , описывающему некоторый объект из O (не входящий в обучающее множество O_1, \dots, O_m), определить этот класс.

Под признаком в контексте раздела будем понимать некоторое обобщенное его значение: конъюнкция (по М.М. Бонгарду), тест [55], эталонный элемент, число градаций при квантовании сигнала (гл. 5) и т.д.

Без ограничения общности, будем считать признаки z_i , $i = \overline{1, n}$ ранжированными, например, по убыванию весовых коэффициентов признаков, вычисленных по какому-либо методу, обзор которых сделан в гл. 4 для разного вида исходной информации и являющихся числовой мерой качества признаков.

²⁸ Лауреат Нобелевской премии 1965г, «Фейнмановские лекции по физике»

Пусть неравенство $|x_j^i - a_j^i| < \delta^i$ означает, что признак z_j корректно распознает i -й объект с точностью δ^i (допустимое расхождение между значениями признака для i -го объекта). Здесь x_j^i, a_j^i – «текущее» значение и «эталонное» (среднее значение признака) z_j для i -го объекта, соответственно). Обозначим через $P(|x_j^i - a_j^i| < \delta^i) = p_{ij}$ вероятность этого события.

Обозначим $\xi_j = \xi_j(m)$ – случайная величина (с.в.), означающая число корректно распознанных объектов признаком $z_j, j = \overline{1, n}$,
 $\xi_j = \sum_{i=1}^m \chi(|x_j^i - a_j^i| < \delta^i)$, где $\chi(x) = \begin{cases} 1, & x = true, \\ 0, & x = false; \end{cases}$; $S_n = \sum_{j=1}^n w_j \xi_j / nm$ – нормированное число распознанных объектов всеми признаками, w_j – весовой коэффициент признака.

6.1. Постановка задачи оценки качества процедур распознавания состояний ДО

Введем с.в. как момент остановки случайного процесса Φ_n :

$$\tau = \inf\{n \geq n_0 : S_n > q_0\} = \inf\{n \geq n_0 : \Phi_n > g_n\}, \quad (6.1)$$

где $g_n = nmq_0 - \sum_{j=1}^n w_j v_j$, q_0 – заданный уровень надежности распознавания,

$\Phi_n = \sum_{j=1}^n \varsigma_j$, $\varsigma_j = w_j (\xi_j - v_j)$, v_i – математическое ожидание с.в. ξ_i , $v_i = M\xi_i$.

Поставим задачу оценить вероятностное распределение величины τ , или минимально необходимого числа признаков, обеспечивающего заданный уровень надежности распознавания, если известно вероятностное распределение признаков (что естественно для приложений, где накоплены соответствующие статистические данные).

Знание распределения величины τ , являющейся пороговым значением качества алгоритма распознавания, позволит: а) оценить требуемое количество признаков (n) для достижения заданного уровня качества распознавания; б) оптимизировать процедуру распознавания по параметрам n, m , а также их весовым коэффициентам. Так, в приложениях (геологических, медицинских и пр.) актуально решение вопроса о соотношении длины выборки данных (m) (например, дорогостоящее бурение дополнительной скважины) и числа характеристических признаков (n) (введение менее дорогостоящего дополнительного анализа для уже имеющихся данных).

6.2. Вероятностная модель пересечения случайным процессом определенного уровня

Для решения поставленной задачи будем использовать вероятностную модель пересечения случайным процессом определенного уровня. Подход, который будет использован для получения оценок вероятностного распределения величины τ , применен А.А. Новиковым [123, 124] для анализа асимптотики вероятности непересечения криволинейной границы суммой независимых разно-распределенных с.в. Суть этого подхода заключается во введении новой вероятностной меры, относительно которой момент пересечения криволинейной границы суммой независимых разнораспределенных с.в. является моментом пересечения мартингалом некоторого постоянного уровня, для которого уже нетрудно получить непосредственно нужные оценки. Задача, таким образом, сводится к доказательству существования подходящей новой вероятностной меры.

Необходимость отдельного изучения этого вопроса связана с тем, что в [123, 163] при рассмотрении моментов останова $\sigma = \inf\{n \geq m : |\Phi_n - f_n| > g_n\}$,

где $\Phi_n = \sum_{i=1}^n \xi_i$, $\{\xi_i, i \geq 1\}$ – последовательность независимых с.в., f_n, g_n – неслучайные функции (границы). При этом накладываемые ограничения на поведение границ f_n, g_n делают невозможным напрямую использовать полученные результаты. Покажем, что, тем не менее, рассуждения, используемые в [123, 163], по существу сохраняются и позволяют получить требуемое преобразование.

Случайные величины ζ_i ограничены (в силу ограниченности с.в. ξ_i , $|\xi_i| \leq m$), поэтому для них выполнено требуемое в [7] условие конечности для всех $|\lambda| < \infty$ функции $M \exp(\lambda \zeta_i)$ и определена функция $\varphi_i(\lambda)$ следующим образом:

$$M \exp(\lambda \zeta_i) = \exp(\varphi_i(\lambda)), \quad (6.2)$$

Последовательность $\{\zeta_i, i \geq 1\}$ будем считать заданной на вероятностном пространстве $\{\Omega, \mathcal{F}, P\}$. Введем последовательность σ -алгебр. Положим $\mathcal{F}_0 = (\emptyset, \Omega)$, $\mathcal{F}_n = \sigma(\zeta_1, \dots, \zeta_n) \in \mathcal{F}$. Пусть (a_i, \mathcal{F}_{i-1}) – предсказуемая последовательность (т.е. для каждого i величины a_i \mathcal{F}_{i-1} -измеримы). Далее на $(\Omega, \mathcal{F}_{n-1})$ вводится вероятностная мера \tilde{P}_n по формуле:

$$\tilde{P}_n(A) = M \chi_A \cdot G_n, \quad A \in \mathcal{F}_n, \quad (6.3)$$

где χ_A – индикатор множества A , а последовательность G_n имеет вид:

$$G_n = \exp \left\{ \sum_{i=1}^n (a_i \zeta_i - \varphi_i(a_i)) \right\}, \quad n \geq 1. \quad (6.4)$$

Можно показать, что последовательность (G_n, \mathcal{F}_n) – мартингал относительно меры P со средним $MG_n = 1$. Обозначим через \tilde{M}_n – математическое ожидание по мере \tilde{P}_n и воспользуемся следующими результатами из [123, 124, 163].

Лемма 6.1. Пусть $\{a_i\}$ ($i = \overline{1, n}$) – последовательность ограниченных предсказуемых с.в.. Тогда

$$\tilde{M}_n((\zeta_i - \varphi'_i(a_i)) / \mathcal{F}_{i-1}) = 0 \quad (\text{п.н.} - \tilde{P}_n), \quad (6.5)$$

$$\tilde{M}_n((\zeta_i - \varphi'_i(a_i))^2 / \mathcal{F}_{i-1}) = \varphi''_i(a_i) \quad (\text{п.н.} - \tilde{P}_n). \quad (6.6)$$

На основании этой леммы можно показать:

а) процесс вида $\mu_i = \sum_{j=1}^i \alpha_j (\zeta_j - \varphi'_j(a_j))$, где $\{a_i\}$ – ограниченная предсказуемая последовательность, является мартингалом относительно $(\mathcal{F}_r, \tilde{P}_r)$ с нулевым средним (это следует из формулы (6.5));

б) возрастающий процесс $\langle \mu_i \rangle$ (т.е. такой предсказуемый процесс, что $\mu_i^2 - \langle \mu_i \rangle$ – мартингал) равен $\langle \mu_i \rangle = \sum_{j=1}^i \alpha_j^2 \varphi''_j(a_j)$ (это следует из формулы (6.6)).

Лемма 6.2. Для $\varphi''_i(\lambda)$ справедливы оценки:

$$c^{-1/2} \sigma_i^2 \leq \varphi''_i(\lambda) \leq c \cdot \sigma_i^2 \quad (6.7)$$

для всех $|\lambda| \leq \ln c \cdot L_i^{-1} / 2$, где $c > 1$, $|\zeta_i| \leq L_i$, $\sigma_i = D\zeta_i$ ($i = \overline{1, n}$).

В контексте поставленной задачи случайная величина ξ_i – есть число распознанных объектов выборки признаком z_i . При заданных вероятностях p_{ji} , $i = \overline{1, n}$, $j = \overline{1, m}$, распределение с.в. ξ_i определяется по формуле:

$$\eta_i(k) = P(\xi_i = k) = \sum_{(j_1, \dots, j_k) \in J_k} \prod_{l=1}^k p_{j_l i} \prod_{t \in \{1, \dots, m\} \setminus J_k} (1 - p_{ti}). \quad (6.8)$$

Здесь J_k – совокупность сочетаний по k объектов из m , $J_k = \{(j_1, \dots, j_l, \dots, j_k), j_l \in (1, \dots, m)\}$ с мощностью множества $|J_k| = C_m^k$.

По определению математическое ожидание и дисперсия с.в. ξ_i равны:

$$v_i = M\xi_i = \sum_{k=1}^m k\eta_i(k), \quad \sigma_i^2 = D\xi_i = \sum_{k=0}^m (k - v_i)^2 \eta_i(k).$$

Случайные величины ξ_i ограничены (в силу ограниченности с.в. ξ_i , $|\xi_i| \leq m$) величиной L_i , равной $\max(v_i, (m - v_i))$, поскольку $|\zeta_i| = |\xi_i - v_i| \leq \max(|0 - v_i|, |m - v_i|) = \max(v_i, (m - v_i)) = L_i$.

В данной работе существование последовательности $\{a_i\}$, участвующей в формулах замены меры (6.3), (6.4), доказывается в следующей лемме.

Лемма 6.3. Если выполнены условия:

$$\kappa |q_0 m - v_i| < \varepsilon_c \sigma_i^2 L_i^{-1}, \quad \varepsilon_c = c^{-1} \ln c / 2, \quad i = \overline{1, n}, \quad (6.9)$$

где κ – некоторая положительная величина, $0 < \kappa < 1$, то существует неслучайная последовательность $\{a_i\}$, удовлетворяющая уравнению:

$$|\varphi'_i(a_i)| = \kappa \cdot \Delta g_i, \quad i = \overline{1, n}, \quad (6.10)$$

Замечание. Как показывают результаты численного моделирования, на самом деле условие (6.9) завышенное.

Пример 6.1. Покажем, что если вероятности p_{ij} , $i = \overline{1, n}$, $j = \overline{1, m}$, будут таковы, что $p_{ij} = \gamma_i$ (все объекты равноценны с точки зрения распознавательной способности отдельного признака) и выполнено условие (6.9): $|m(q_0 - \gamma_i)| < \varepsilon_c \cdot \min(\gamma_i, (1 - \gamma_i))$, $\varepsilon_c = \ln c \cdot c^{-1} / 2$, то имеет место утверждение (6.10) леммы 6.3.

В этом случае с.в. ξ_i имеют биномиальное распределение $P(\xi_i = k) = C_m^k \gamma_i^k (1 - \gamma_i)^{m-k}$ с параметрами (m, v_i) , где $v_i = m \cdot \gamma_i$, и дисперсией $D\xi_i = m \cdot \gamma_i \cdot (1 - \gamma_i)$. Подвижная граница в этом случае имеет вид: $g_n = nmq_0 - \sum_{i=1}^n v_i = m(nq_0 - \sum_{i=1}^n \gamma_i)$, при этом для естественного случая, когда $q_0 - \gamma_i > 0$, приращение: $\Delta g_i = g_i - g_{i-1} = m \cdot (q_0 - \gamma_i) \geq 0$. Покажем возможность выполнения (6.10) при соблюдении условий (6.9).

Для выполнения (6.10), как следует из леммы 6.2 и монотонности функции $\varphi'_i(a_i)$, необходимо должно выполняться неравенство:

$$\Delta g_i < \Lambda_i c^{-1} \sigma_i^2 = c^{-1} \ln c L_i^{-1} m \gamma_i (1 - \gamma_i) / 2 = \varepsilon_c \min(\gamma_i, (1 - \gamma_i)), \quad \varepsilon_c = c^{-1} \ln c / 2. \quad (6.11)$$

Найдем величину L_i :

$$|\zeta_i| = |\xi_i - v_i| \leq L_i = \max(|0 - m \cdot \gamma_i|, |m - m \cdot \gamma_i|) = m \cdot \max(\gamma_i, (1 - \gamma_i)),$$

и с учетом (6.11) имеет место (6.10). Для случая, когда $m(q_0 - \gamma_i) - b/n < 0$ утверждение (6.10) леммы 6.3 доказывается аналогично.

Далее введем с.в. $\bar{\zeta}_i$ по формуле: $\bar{\zeta}_i = \zeta_i - \Delta g_i$, $\Delta g_i = g_i - g_{i-1} = m \cdot q_0 - v_i$, полагая $g_0 = mq_0$. Перепишем (6.1), обозначая через $\bar{\Phi}_n$ сумму $\bar{\Phi}_n = \sum_{i=1}^n \bar{\zeta}_i$ (имея в виду, что $\sum_{i=1}^n \Delta g_i = g_n - g_0$):

$$\begin{aligned} \tau &= \inf\{n \geq 1: \sum_{i=1}^n \zeta_i > g_n\} = \inf\{n \geq 1: \bar{\Phi}_n + \sum_{i=1}^n \Delta g_i > g_n\} = \inf\{n \geq 1: \bar{\Phi}_n > g_0\} = \\ &= \inf\{n \geq 1: \bar{\Phi}_n > mq_0\}. \end{aligned} \quad (6.12)$$

Лемма 6.4. Последовательность $\{\bar{\Phi}_i, i = \overline{1, n}\}$ является мартингалом относительно $(\mathcal{F}_n, \tilde{P}_n)$ с нулевым средним.

Таким образом, относительно новой меры \tilde{P}_n за момент достижения требуемого качества процедуры распознавания, являющегося функцией числа используемых признаков, объема обучающей выборки и вероятностей p_{ij} , $i = \overline{1, n}$, $j = \overline{1, m}$ может быть принят момент первого пересечения снизу мартингалом $\{\bar{\Phi}_n, \mathcal{F}_n\}$ постоянного уровня.

Вывод оценки для вероятности обеспечения требуемого качества процедуры распознавания существенно опирается на лемму Блэкуэлла и Фридмана (см., например, [7]) о поведении хвостов распределений моментов первого выхода из единичной полосы мартингалов с ограниченными скачками. Заметим, что для получения нижних оценок величины $P(\tau > n)$ могут быть применены также лемма А.А. Новикова [123] и лемма А.Н. Ширяева [163] о моменте первого выхода мартингала за отрицательный уровень. Кратко прокомментируем возможность применения каждого из перечисленных результатов.

1. Для применения леммы Блэкуэлла и Фридмана требуется ограниченность скачков мартингала $\{\bar{\Phi}_n, \mathcal{F}_n\}$, что будет показано ниже.

2. Условия леммы А.А. Новикова предполагают ограниченность квадратов скачков мартингала $\{\bar{\Phi}_n, \mathcal{F}_n\}$: $\tilde{M}_n \left[(\Delta \bar{\Phi}_n)^+ \right]^2 \leq d_1 < \infty$, $x^+ = \max(x, 0)$. Данное требование выполнимо по мере \tilde{P}_n .

3. Условия леммы А.Н. Ширяева предполагают квадратичную интегрируемость мартингала $\{\bar{\Phi}_n, \mathcal{F}_n\}$, в справедливости которого не сложно убедиться.

6.3. Двусторонние оценки для вероятности достижения гарантированного уровня качества процедуры распознавания

Общая схема получения основного результата заключается в следующем.

1. Определяется функция $\varphi_i(\lambda)$: $M \exp(\lambda \zeta_i) = \exp(\varphi_i(\lambda))$. Это возможно в силу ограниченности с.в. ζ_i , $|\zeta_i| \leq m$. Функция $\varphi_i(\lambda)$ является бесконечно дифференцируемой, и непосредственно из ее определения следует, что $\varphi_i'(\lambda) = M(\zeta_i \exp\{\lambda \zeta_i - \varphi_i(\lambda)\})$, $\varphi_i''(\lambda) = M\left(\left(\zeta_i - \varphi_i'(\lambda)\right)^2 \exp\{\lambda \zeta_i - \varphi_i(\lambda)\}\right)$.

2. На основе функций $\varphi_j(\lambda)$ ($j = \overline{1, m}$) на $(\Omega, \mathcal{F}_{n-1})$ вводится вероятностная мера \tilde{P}_n по формуле: $\tilde{P}_n(A) = M \chi_A \cdot G_n$, $A \in \mathcal{F}_n$, где χ_A – индикатор множества A , а последовательность G_n имеет вид:

$G_n = \exp\left\{\sum_{j=1}^n (a_j \zeta_j - \varphi_j(a_j))\right\}$, $n \geq n_0$, где $\{a_i\}$ ($i = \overline{1, n}$) – последовательность ограниченных предсказуемых с.в., подлежащих определению, при этом существенно используется факт принципиальной разрешимости уравнения $\varphi_j'(a_j) = x$ при малых $x \in R$ (лемма 6.3). Известно [163], что последовательность (G_n, \mathcal{F}_n) – мартингал относительно меры P со средним $M G_n = 1$.

3. Доказывается, что относительно $(\mathcal{F}_n, \tilde{P}_n)$ исходный процесс Φ_n является мартингалом с нулевым средним, а подвижная граница g_n – некоторым постоянным уровнем, для которого уже нетрудно получить непосредственно нужные оценки (на основании лемм Блэкуэлла и Фридмана, А.А. Новикова, А.Н. Ширяева).

4. Осуществляется корректный возврат в исходное пространство $\{\Omega, \mathcal{F}, P\}$.

Сформулируем основной результат проведенного исследования.

Теорема 6.1. Пусть признак z_i , $i = \overline{1, n}$, распознает объект $O_j = \{a_1^j, \dots, a_n^j\}$, $j = \overline{1, m}$, с вероятностью p_{ji} , и справедливы условия леммы 3. Тогда имеют место оценки для надежности распознавания $R_n = P(\tau > n)$ (вероятности обеспечения заданного качества распознавания):

$$\begin{aligned} R_n &\leq c_1 \exp\left\{-c_2 (q_0 m)^{-2} \left[\sum_{j=1}^n \sigma_j^2 c^{-1/2}\right] + c_3 \sum_{i=1}^n \nu_i\right\}, \\ R_n &\geq \exp\left\{-c_4 \left[\alpha^2 + (q_0 m)^{-2} c \sum_{j=1}^n \sigma_j^2\right] - c_5 \sum_{i=1}^n \nu_i\right\}, \end{aligned} \quad (6.13)$$

где $v_i = (q_0 m - v_i) / \sigma_i^2$, величины $c_i, i = \overline{1,5}$, не зависят от n и являются известными функциями от величин (α, p, q, c) : $c_1 = (5 / 2\alpha)^{1/2 p}$, $c_2 = \pi^2 (1 - 3\alpha^{1/2}) c^{-1/2} / (8p)$, $c_3 = (q - 1)c^2$, $c_4 = \pi^2 p (1 + 15\alpha^{3/4}) / 8$, $c_5 = c^2 \frac{q+1}{q-1}$, $\alpha \leq 1/16, 1/p + 1/q = 1, p > 1, c > 1$.

Прокомментируем утверждение теоремы 6.1.

В интеллектуальных системах, основанных на методах тестового распознавания образов, для принятия решений используются тесты (минимальные наборы признаков, гл. 4), с определенной вероятностью разделяющие обучающий материал на классы. Пусть ξ_i – случайная величина (с.в.), означающая число корректно распознанных объектов на контрольной выборке, совершенных тестом $T_i = \{z_{l_1}, \dots, z_{l_{g_i}}\}$, $g_i \leq n, l_k \in \{1, \dots, n\}$, n – число признаков, m – число

объектов. Тогда с.в. $\Phi_n = \sum_{i=1}^n \xi_i / nm$ представляет собой долю корректно распо-

знанных объектов n первыми тестами. С ростом числа используемых тестов точность решающего правила увеличивается, т.к. итоговое решение принимается процедурой голосования «по большинству». Однако, во-первых, использование всех возможных тестов вряд ли возможно даже при современных вычислительных возможностях; во-вторых, при измерении признаков возможны погрешности, которые могут вносить нежелательные «коррекции», усиливающиеся при увеличении числа используемых при принятии решений тестов (так называемое явление переобучения [47]).

Применение полученных оценок для вероятности момента первого достижения гарантированного уровня надежности распознавания (по числу используемых тестов) потребует знания вероятностных характеристик с.в. ξ_i , в частности, вероятность η_{ik} того, что тест $T_i = \{z_{l_1}, \dots, z_{l_{g_i}}\}$ корректно распознает k

объектов: $\eta_{ik} = P(\xi_i = k) = \sum_{(j_1, \dots, j_s, \dots, j_k) \in J_k} \prod_{s=1}^k \prod_{t=1}^{g_i} r_{j_s l_t} \prod_{j_u \notin J_k} \prod_{l_d \in Z \setminus T_i} (1 - r_{j_u l_d})$. Здесь J_k

– совокупность сочетаний по k объектов из m ,

$J_k = \{(j_1, \dots, j_l, \dots, j_k), j_l \in (1, \dots, m)\}$ с мощностью множества $|J_k| = C_m^k$;

$Z = \{z_1, \dots, z_n\}$ – множество всех признаков; r_{ij} – вероятность корректного распо-

знавания объекта O_j тестом $T_i, i \geq 1, j = \overline{1, m}$. Если под корректным распознаванием объекта тестом [52] понимать корректное распознавание этого объекта

всеми (независимыми) признаками, входящих в тест, то $r_{ji} = \prod_{t=1}^{g_i} p_{j l_t}$.

Выводы. Для каждой задачи распознавания (состояний ДО) существует некоторый максимально возможный уровень правильного распознавания, что с практической точки зрения прозрачно интерпретируется: недостаточная исходная информативность признаков не может гарантировать получение высокого качества классификации объектов из генеральной совокупности за счет алгоритмических построений. Весьма обширные экспериментальные исследования (в том числе и в этой книге, см. гл. 2, 4, 5, 8) свидетельствуют, что использование разных алгоритмов, реализующих принципиально отличающиеся подходы, а также коллектив алгоритмов приводят к примерно одному и тому же уровню ошибок для конкретного типа ВР. Это является косвенным подтверждением того, что, во-первых, определяющим фактором при распознавании состояний ДО является информативность признакового пространства, а не тип используемых алгоритмов; во-вторых, для любой модели ДО можно оценить примерный уровень безошибочного распознавания состояний ДО, выше которого достичь принципиально невозможно независимо от числа и типа используемых алгоритмов (по крайней мере, не используя дополнительных предположений).

В главе 6 рассмотрен вероятностный подход для выяснения условий, обеспечивающих получение двусторонней оценки гарантированного качества статистической процедуры распознавания с целью оптимального соотношения «число признаков – число объектов», который может быть использован при выборе алгоритма распознавания и настройке определенного алгоритма по обучающей выборке статистического характера. Дальнейшее исследование этого вопроса может быть связано с выяснением статистических свойств предложенных оценок, а также со сравнением нижних оценок, полученных с помощью вышеупомянутых ранее полученных результатов в работах [123, 124].

Результаты, представленные в 6-й главе, опубликованы в работах [73, 74, 89, 138].

«Разнообразие управляющей системы должно быть не меньше разнообразия управляемого объекта»

/Закон необходимого разнообразия У. Эшби/

«Правила запрета гласят: бессмысленно тратить энергию и время на насилие над сложными системами, надо знать, как они функционируют, с минимальными усилиями возбуждать то, что годится для них, учитывая собственные реакции объекта на внешние воздействия» /А.А. Самарский²⁹/

Глава 7. АДАПТИВНОЕ УПРАВЛЕНИЕ СЛАБОФОРМАЛИЗОВАННЫМ ДИНАМИЧЕСКИМ ОБЪЕКТОМ НА ОСНОВЕ МОДЕЛИ СОСТОЯНИЯ

Система управления на основе традиционных принципов работает только с известными и отраженными в математической модели закономерностями связей параметров. В настоящее время актуальны АСУ на основе баз данных с активными многопараметрическими и слабоструктурированными объектами управления. Традиционные системы управления в этом случае практически не эффективны [2, 7, 14, 105, 106, 149, 158]. Одним из способов частичного решения данной проблемы является применение методов искусственного интеллекта (распознавания образов как идентификации состояния некоторого объекта), методов поддержки принятия решений и теории информации при построении систем идентификации и прогнозирования состояний среды при управлении слабоструктурированными объектами.

Вопрос о применении систем распознавания образов для принятия решений об управляющем воздействии не нов, но конструктивный общий подход к формированию управления отсутствует по понятной причине: задача сравнения и выбора наилучших вариантов управления, получаемых на основе различных сочетаний управляющих факторов, по определенным критериям чрезвычайно непростая и *NP*-полна.

В данном разделе по заданному целевому состоянию объекта управления и на основе распознавания состояний объекта управления (как классов) построен регулятор, учитывающий влияния признаков, способствующих или препятствующих переходу объекта в определенное состояние, и, на этой основе, вырабатывающий решение об управляющем воздействии. Задача выбора признаков по состоянию является обратной задачей прогнозирования, т.к. при прогнозировании, наоборот, определяется состояние по признакам. При этом признаки могут иметь разную природу происхождения: 1) характеризовать предысторию объекта управления и его актуальное состояние управления; 2) управляющие факто-

²⁹ Выдающийся российский математик, академик РАН (1919-2008).

ры; 3) внешние воздействия (в т.ч. шумовые). Блок распознавания состояний и признаков, способствующих их наступлению, участвует, во-первых, в идентификации состояния объекта управления; во-вторых, в выработке управляющих воздействий.

7.1. Постановка задачи синтеза управления по методу АКАР

Пусть ДО задан в виде:

$$\begin{aligned}\dot{\mathbf{x}} &= f(\mathbf{x}, \boldsymbol{\theta}, \mathbf{u}, t), \\ \mathbf{y} &= h(\mathbf{x}, t),\end{aligned}\tag{7.1}$$

$\mathbf{x} \in R^n$ - вектор состояний, $\boldsymbol{\theta} \in R^k$, - вектор параметров и возмущающих воздействий, $\mathbf{u} \in R^m$, $m < n$, - вектор управления, $f(\cdot)$, $h(\cdot)$ – непрерывные нелинейные вектор-функции, $\mathbf{y} \in R^m$, - вектор выходных переменных (измерений).

Ставится задача нахождения закона управления в виде

$$\mathbf{u}(\boldsymbol{\psi}) = \mathbf{u}(\mathbf{x}),\tag{7.2}$$

обеспечивающего перевод объекта управления (7.1) из произвольного начального состояния \mathbf{x}_0 в некоторой области фазового пространства в заданное состояние и его стабилизации в некоторой окрестности инвариантного многообразия (аттрактора):

$$\boldsymbol{\psi}(\mathbf{x}) = 0,\tag{7.3}$$

или другими словами: попадание в окрестность инвариантного многообразия, а затем дальнейшее асимптотически устойчивое движение изображающей точки (ИТ) вдоль этого многообразия в желаемое состояние ДО (задаваемое целевое множество, в частности, в начало координат).

Макропеременные $\boldsymbol{\psi}(\mathbf{x})$ должны удовлетворять основному функциональному уравнению [70]:

$$\omega \dot{\boldsymbol{\psi}}(t) + \varphi(\boldsymbol{\psi}) = 0,\tag{7.4}$$

которое в условиях: $\omega > 0$, $\varphi(\mathbf{x})$ – однозначная, непрерывная и дифференцируемая, $\varphi(0) = 0$, $\varphi(\mathbf{x}) \cdot \mathbf{x} > 0$ при любых $\mathbf{x} \neq 0$, является уравнением Эйлера-Лагранжа относительно устойчивых экстремалей, доставляющих минимум на траекториях движения замкнутой системы сопровождающему оптимизирующему функционалу (СОФ) вида:

$$\Phi = \int_0^{\infty} (\omega^2 \dot{\boldsymbol{\psi}}^2(t) + \varphi^2(\boldsymbol{\psi})) dt.\tag{7.5}$$

Весовой коэффициент (ВК) ω содержательно определяет задаваемое время движения изображающей точки системы до пересечения многообразий и является параметром настройки регулятора.

Определим производную функции $\psi_i(t)$: $\dot{\psi}_i(t) = \sum_{k=1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_k} \dot{x}_k$ и подставим в СОФ с учетом исходной СДУ:

$$\begin{aligned} \dot{x}_j(t) &= f_j(y_1, \dots, y_n), \quad j = \overline{1, n-m}, \\ \dot{x}_j(t) &= f_j(y_1, \dots, y_n) + u_j, \quad j = \overline{n-m+1, n}. \end{aligned} \quad (7.6)$$

Получим:

$$\begin{aligned} \dot{\psi}_i(t) &= \sum_{j=1}^{n-m} \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} f_j(y_1, \dots, y_n) + \sum_{j=n-m+1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} (f_j(y_1, \dots, y_n) + u_j) = \\ &= \sum_{j=1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} f_j(y_1, \dots, y_n) + \sum_{j=n-m+1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} u_j, \\ \Phi &= \int_0^{\infty} \sum_{i=1}^m \left[\phi_i^2(\psi_i) + \omega_i^2 \left(\sum_{j=1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} f_j(y_1, \dots, y_n) + \sum_{j=n-m+1}^n \frac{\partial \psi_i(x_1, \dots, x_n)}{\partial x_j} u_j \right)^2 \right] dt. \end{aligned}$$

Такая форма СОФ отражает общие свойства как исходного объекта, так и его системы управления.

Решение задачи адаптивного управления по АКАР [70] и методу I&I [2] с привлечением синергетического подхода осуществимо при выполнении ряда следующих ограничительных условий:

1. Существование глобально устойчивой целевой системы, удовлетворяющей технологическим целям.
2. Существование целевого инвариантного многообразия по отношению к исходной системе уравнений объекта.
3. Многообразие должно иметь возможность «неявного» задания (локальное целевое условие) в виде $\psi(x)=0$, где x – состояние объекта, а $\psi(x)=0$ описывает состояние целевой системы.
4. Многообразие должно быть аттрактивным и все решения исходной системы должны быть ограниченными.
5. Стабилизируемость исходной модели состояния объекта.

Таким образом, для применения синергетических методов управления I&I и АКАР чрезвычайно важна информация о структуре фазового пространства, в частности, информация о существовании инвариантных многообразий, их аналитическом описании и времени их достижения (время вхождения в аттрактор), т.к. именно так интерпретируются неизвестные весовые коэффициенты функционала качества (см. пример 7.3; п. 9.1). Отметим, что вопрос о времени вхождения в аттрактор в АКАР является открытым, и подлежит в каждой задаче определению. Один из приемов определения этого параметра изложен в п. 9.2.

Будем далее рассматривать идентификацию ДО (идентификацию состояний ДО) как один из этапов проектирования регуляторов.

7.2. Постановка задачи синтеза управления ДО в пространстве состояний

Наиболее популярной и имеющей простое геометрическое толкование является следующая форма модели ДО:

$$\begin{aligned} \dot{x}_1(t) &= F_1(x_1, \dots, x_n), \\ &\dots\dots\dots \\ \dot{x}_n(t) &= F_n(x_1, \dots, x_n), \end{aligned} \tag{7.7}$$

к которой можно привести любую систему ОДУ, разрешенную относительно старших производных (даже если она содержит явную зависимость от времени). Систему (7.7) путем замены переменных можно привести к виду [17, 18]:

$$\begin{aligned} \dot{y}_1(t) &= y_2, \\ \dot{y}_2(t) &= y_3, \\ &\dots\dots\dots \\ \dot{y}_D(t) &= F(y_1, \dots, y_D), \end{aligned} \tag{7.8}$$

Рассмотрим класс динамических объектов, имеющих описание:

$$\begin{aligned} \dot{x}_1(t) &= x_2, \\ &\dots\dots\dots \\ \dot{x}_{n-1}(t) &= x_n, \\ \dot{x}_n(t) &= f(x, \theta, \xi, u). \end{aligned} \tag{7.9}$$

$x(t) = (x_1, \dots, x_n) \in R^n$ - вектор переменных состояния; $\theta(t) \in R^m$ - вектор параметров; $u(t)$ - управляющий сигнал; $\xi(t) \in R$ - возмущения, $f(\cdot)$ - неизвестная нелинейная (в общем виде) функция.

Согласно [70] будем полагать, что управление осуществляется в пространстве состояний объекта с использованием макропеременной состояния $\psi(x)$, некоторой произвольной дифференцируемой или кусочно-непрерывной функции, обращение в нуль которых задает инвариантные многообразия ДО, отражающих целевое макросостояние ДО.

Требуется сформировать управляющие воздействия, переводящие объект из заданного начального состояния x_0 в окрестность $\psi(x) = 0$ и доставляющие минимум функционалу вида (с учетом одномерности управления):

$$J = \int_0^{\infty} (\psi^2(x) + \omega \dot{\psi}^2(x)) dt, \tag{7.10}$$

где весовой коэффициент ω содержательно определяет задаваемое время движения изображающей точки системы до пересечения многообразий и является параметром настройки регулятора; $\psi(x)$ – непрерывная функция. Здесь функцию $\varphi(\psi)$, определенную в (7.4), возьмем в виде: $\varphi(\psi)=\psi$.

7.3. Регулятор для управления слабоформализованным ДО в пространстве состояний

Для ряда задач стохастической теории управления справедлив принцип (теорема) разделения [160]. Основное содержание утверждения о разделении состоит в том, что оптимальная операция, разрешающая задачу о построении управления, действующего на основе неполных или недостоверных измерений, может быть представлена в виде суперпозиции решений двух частных задач. Первая из них представляет собой задачу управления для исходной системы, но с полным составом наблюдения, а вторая - задачу оценивания состояния и/или компонент траектории системы при фиксированном управлении (рис. 7.1).



Рис. 7.1. Структурная схема совмещения традиционного управления по АКАР с адаптацией по состоянию для сложных ДО с неопределенностью

В данном разделе построен регулятор, совмещающий аналитическое управление АКАР с адаптацией по состоянию, как множеству возможных значений фазовой траектории ДО (рис. 7.1).

Обозначим $u_i = \begin{cases} v_i, & \mathbf{x} \in \Omega_i \\ 0, & \mathbf{x} \notin \Omega_i \end{cases}$, Ω_i – i -е состояние (подмножество значений характеристических признаков состояния), $i = \overline{1, I}$.

Теорема 7.1. Для класса динамических объектов, описываемых моделью (7.7) с вектором состояний $x \in R^n$, вектором управления $u \in R^1$, $\xi \in R^1$ - вектором возмущений, $f \in R^1$ - неизвестной (нелинейной) вектор-функции, представимой в виде: а) $f_1(\mathbf{x}, \boldsymbol{\theta}, \xi) + f_2(u)$, или б) $f_1(\mathbf{x}, \boldsymbol{\theta}, \xi) f_2(u)$, существует закон управления в виде:

$$u(\psi) = u(\mathbf{x}) = \nu^T u, \quad \nu = \|v_{li}\|_{I \times 1}, \quad u = \|u_{li}\|_{I \times 1}, \quad (7.11)$$

где u_i, v_i – управляющее воздействие и весовой коэффициент управляющего воздействия в i -м состоянии динамического объекта (7.9), соответственно, удовлетворяющие соотношениям:

$$u_i = \begin{cases} u_i' : \dot{u}_i' = -\gamma v_i \frac{\partial \psi}{\partial x_n}, & \mathbf{x} \in \Omega_i, \frac{\partial \psi}{\partial x_n} \neq 0, i = \overline{1, I}, \\ 0, & \mathbf{x} \notin \Omega_i \end{cases} \quad (7.12)$$

который

1) обеспечивает перевод объекта управления (7.9) из произвольного начального состояния $x_0 \in R^n$ в некоторой области фазового пространства в заданное состояние $\psi(x) = 0$ и его стабилизацию в некоторой окрестности инвариантного многообразия $\psi(x) = 0$ с макропеременной $\psi(x)$, ограниченной условием $\frac{\partial \psi}{\partial x_n} \neq 0$, для класса систем (7.9) с правой частью типа а); ограниченной условием

$\frac{\partial \psi}{\partial x_n} \neq 0, f \neq 0$, для класса систем (7.9) с правой частью типа б);

2) доставляет минимум на траекториях движения замкнутой системы сопровождающему оптимизирующему функционалу вида (7.10);

3) обеспечивает асимптотическую устойчивость объекта (7.9).

На рис. 7.2 размещена предлагаемая функциональная блок-схема адаптивного управления сложным объектом типа (7.9), где использованы обозначения и сокращения: $i(t)$ – метка (номер) состояния ДО, $\xi(t)$ – неопределенность в описании ДО (шум), $x(t)$ – многомерная характеристика ДО во времени, $u(i, t)$ – управляющее воздействие, зависящее от информации о состоянии ДО, $\nu(i, t)$ – вектор параметров оптимизации регулятора.

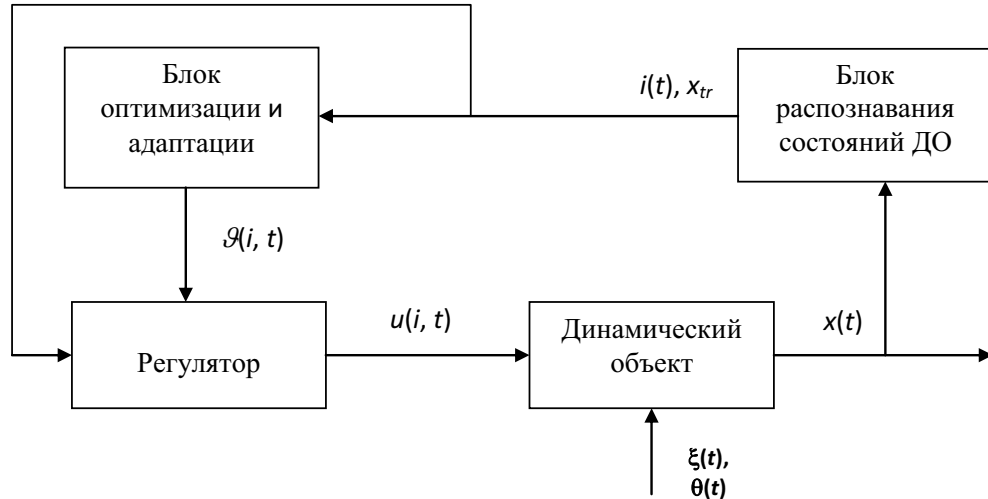


Рис. 7.2. Схема управления слабоформализованным ДО

Изложим поэтапно процедуру построения регулятора (рис.7.2) поэтапно для случая а) и покажем затем изменения в рассуждениях для справедливости утверждения теоремы 7.1 и для случая б).

Первый этап. На этом этапе используется метод АКАР с полной информацией. Исходя из заданного критерия качества работы системы (7.10), учитывая назначенные целевые ограничения $\psi(x) = 0$ и априорную информацию о числе состояний системы (7.9), производится синтез основной структуры системы управления.

Следуя АКАР, макропеременная ψ должна удовлетворять основному функциональному уравнению:

$$\omega \dot{\psi} + \psi = 0, \quad (7.13)$$

которое при $\omega > 0$ является уравнением Эйлера-Лагранжа относительно устойчивых экстремалей, доставляющих минимум на траекториях движения замкнутой системы сопровождающему оптимизирующему функционалу (7.10). Учитывая,

что $\dot{\psi} = \sum_{k=1}^n \frac{\partial \psi}{\partial x_k} \dot{x}_k = \sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} (f_1 + f_2(u))$ и подставляя это соотношение в (7.13), получим:

$\omega \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} (f_1 + f_2(u)) \right) + \psi = 0$. Отсюда следует:

$$f_2(u_c) = - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} f_1 + \omega^{-1} \psi \right), \quad \frac{\partial \psi}{\partial x_n} \neq 0. \quad (7.14)$$

или

$$u_c = (f_2)^{-1} \left(- \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} f_1 + \omega^{-1} \psi \right) \right), \quad \frac{\partial \psi}{\partial x_n} \neq 0, \quad (7.15)$$

где $(f_2)^{-1}$ - обратная функция для f_2 .

При полной определенности относительно функции f синтез управления в пространстве состояний по методу АКАР (аналогично и по методу I&I, или «обход интегратора») закончен.

Второй этап. Управляющий сигнал будем выбирать по правилу:

$$u = u_i \in U_i, \text{ если } x(t) \in \Omega_i \quad (7.16)$$

где численное значение управляющего сигнала $u(t) = u_i$ для каждого $t \in (t_i, t_{i+1})$ подлежит определению (U_i - множество допустимых управляющих воздействий в состоянии Ω_i) и должно (при решении задачи синтеза) доставлять оптимум критерию (7.10), и адаптивно учитывать особенности $x(t)$, устойчиво выводя объект на заданное многообразие

$$\psi(x) = 0.$$

Будем искать оптимальный управляющий сигнал u в виде:

$$f_2(u) = \sum_{i=1}^l \nu_i u_i \quad (u = u^T \nu, u^T = (u_1, \dots, u_l), \nu^T = (\nu_1, \dots, \nu_l)),$$

где ν_i - весовой коэффициент i -го управляющего воздействия u_i , подлежащего определению.

Представим последнее уравнение в (7.9) в виде:

$$\dot{x}_n(t) = f_1 + f_2(\hat{u}) + (f_2(u) - f_2(\hat{u})) \quad (7.17)$$

где в качестве аргумента второго слагаемого $f_2(\hat{u})$ возьмем выражение для управления $f_2(u_c)$ (7.6), полагая $\hat{u} = u_c$, а в качестве третьего слагаемого выражение:

$$f_2(u) - f_2(\hat{u}) = \sum_{i=1}^l \nu_i (u_i - \hat{u}_i),$$

аппроксимируя $f_2(u_c)$ выражением $\sum_{i=1}^l \nu_i \hat{u}_i$, где ν_i - весовой коэффициент i -го

управляющего воздействия u_i . Величины ν_i, u_i , подлежат дальнейшему определению. Последнее уравнение в системе (7.9) при таком преобразовании примет вид:

$$\begin{aligned} \dot{x}_n(t) = \dot{x}_n(t) &= f_1 + f_2(\hat{u}) + (f_2(u) - f_2(\hat{u})) = f_1 - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} f_1 + \omega^{-1} \psi \right) + \\ &+ \sum_{i=1}^l \nu_i (u_i - \hat{u}_i) = - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) + \sum_{i=1}^l \nu_i (u_i - \hat{u}_i). \end{aligned}$$

Перепишем систему (7.9) с учетом полученного выражения:

$$\begin{aligned} \dot{x}_1(t) &= x_2, \\ &\dots\dots\dots \\ \dot{x}_{n-1}(t) &= x_n, \\ \dot{x}_n(t) &= - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) + \sum_{i=1}^l \nu_i (u_i - \hat{u}_i). \end{aligned} \tag{7.18}$$

Таким образом, в силу отсутствия полной информации об объекте (7.9) и ограничительных условий на управление система управления синтезирована с точностью до параметров управления, однако ее оптимальные параметры с помощью аналитических процедур отыскать невозможно: результатом второго этапа является только структура системы управления (7.18).

Третий этап. На третьем этапе конструирования регулятора для слабоформализованного динамического объекта следует произвести выбор таких функций (u_1, \dots, u_l) , при которых будет обеспечиваться асимптотический перевод системы управления из периферийных значений функционалов качества в их минимальное состояние. Такой подход позволяет использовать для поиска указанных вектор-функций конструктивный аппарат функций Ляпунова.

В соответствии с видом системы (7.18) и теоремы Ляпунова о достаточном условии устойчивости решений выберем функцию Ляпунова в виде:

$$V(t) = \frac{1}{2} \left[\psi^2 + \gamma^{-1} \sum_{i=1}^l \nu_i (u_i - \hat{u}_i)^2 \right], \gamma > 0. \tag{7.19}$$

Производная функции Ляпунова:

$$\dot{V}(t) = \psi \dot{\psi} + \gamma^{-1} \sum_{i=1}^l \nu_i (u_i - \hat{u}_i) (\dot{u}_i - \dot{\hat{u}}_i) = \psi \dot{\psi} + \gamma^{-1} \sum_{i=1}^l \nu_i (u_i - \hat{u}_i) \dot{u}_i. \tag{7.20}$$

Получим соотношение для $\dot{\psi}$ на основе (7.9), (7.18):

Выражение для $\dot{\psi}$:

$$\begin{aligned}\dot{\psi} &= \sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} \left(f_1 - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \frac{\partial \psi}{\partial x_n} f_1 + \omega^{-1} \psi \right) + \sum_{i=1}^I \nu_i (u_i - \hat{u}_i) \right) = \\ &= -\omega^{-1} \psi + \frac{\partial \psi}{\partial x_n} \sum_{i=1}^I \nu_i (u_i - \hat{u}_i).\end{aligned}$$

Подставим в (7.20) выражение для $\dot{\psi}$:

$$\begin{aligned}\dot{V}(t) &= \psi \dot{\psi} + \gamma^{-1} \sum_{i=1}^I (u_i - \hat{u}_i) \dot{u}_i = -\omega^{-1} \psi^2 + \psi \frac{\partial \Psi}{\partial x_n} \sum_{i=1}^I (u_i - \hat{u}_i) + \\ &+ \gamma^{-1} \sum_{i=1}^I (u_i - \hat{u}_i) \dot{u}_i = -\omega^{-1} \psi^2 + \sum_{i=1}^I (u_i - \hat{u}_i) \left(\nu_i \psi \frac{\partial \psi}{\partial x_n} + \gamma^{-1} \dot{u}_i \right).\end{aligned}$$

Требование устойчивости решения (7.18) по Ляпунову ($\dot{V} < 0$) обеспечивается неравенством: $\sum_{i=1}^I (u_i - \hat{u}_i) \left(\nu_i \psi \frac{\partial \psi}{\partial x_n} + \gamma^{-1} \dot{u}_i \right) < \omega^{-1} \psi^2$, которое при выполнении условия (7.12) теоремы 7.1 имеет место.

Таким образом, в соответствии с доказанным регулятор с точностью до весовых коэффициентов управляющих воздействий имеет вид ($\eta_i = (u_i - \hat{u}_i)$):

$$\begin{aligned}\dot{x}_1(t) &= x_2, \\ &\dots\dots\dots \\ \dot{x}_{n-1}(t) &= x_n, \\ \dot{x}_n(t) &= - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) + \sum_{i=1}^I \nu_i \eta_i, \\ \dot{\eta}_i &= -\gamma \nu_i \psi \frac{\partial \psi}{\partial x_n}, \quad i = \overline{1, I}.\end{aligned}\tag{7.21}$$

Четвертый этап. На четвертом этапе формирования регулятора для слабоформализованного ДО осуществляется конструктивный выбор весовых функций ν_1, \dots, ν_I , удовлетворяющих условию (7.12).

Правило для выбор величин ν_1, \dots, ν_I зависит от требований к функционированию системы (7.9), (7.21) и уровня априорной информации.

Пусть экспертно заданы целевые траектории $x_{tr}(t)$ для части координат, удовлетворяющие целевому многообразию $\psi(x)=0$ ($|\psi(x_{tr}(t))| < \varepsilon, \forall \varepsilon > 0$ при $t \rightarrow \infty$). Тогда выбор коэффициентов $\nu_i, i = \overline{1, I}$, возможен из условия $\dot{x}(t) = \dot{x}_{tr}(t)$. Так для рассматриваемой системы (7.21) при $I=1$ и предполагае-

мой целевой траектории x_{1tr} (модель по измеренным данным) будут справедливы соотношения:

$$\begin{aligned} \dot{x}_{1tr}(t) &= x_2, \\ \dot{x}_2(t) &= x_3 = \ddot{x}_{1tr}(t), \\ &\dots\dots\dots \\ \dot{x}_{n-1}(t) &= x_n = x_{1tr}^{(n-1)}(t), \\ x_{1tr}^{(n)}(t) &= - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) + \nu \eta, \\ \dot{\eta} &= -\gamma \nu \psi \frac{\partial \psi}{\partial x_n}, \end{aligned}$$

из которых следует система для определения управляющего воздействия и его весового коэффициента:

$$\begin{aligned} \nu \eta &= x_{1tr}^{(n)}(t) + \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right), \\ \dot{\eta} &= -\gamma \nu \psi \frac{\partial \psi}{\partial x_n}. \end{aligned}$$

Решая систему, получим соотношение для управляющего воздействия:

$$\dot{\eta} \eta = -\gamma \left[x_{1tr}^{(n)}(t) + \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) \right] \psi \frac{\partial \psi}{\partial x_n}.$$

Изменения в доказательстве для случая б) заключаются в следующем. Правую часть уравнения (7.11) представим в виде:

$$\dot{x}_n(t) = f_1 f_2(\hat{u}) + f_1(f_2(u) - f_2(\hat{u})),$$

где $f_2(u) - f_2(\hat{u}) = \sum_{i=1}^I \tilde{v}_i(u_i - \hat{u}_i)$ и последнее уравнение в системе (7.9) при таком преобразовании примет вид:

$$\begin{aligned} \dot{x}_n(t) &= f_1 f_2(\hat{u}) + f_1(f_2(u) - f_2(\hat{u})) = f_1 \varphi \left(\varphi^{-1} \left(- \left[f_1 \frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) \right) \right) + \\ &+ f_1 \sum_{i=1}^I \tilde{v}_i(u_i - \hat{u}_i) = - \left[\frac{\partial \psi}{\partial x_n} \right]^{-1} \left(\sum_{k=1}^{n-1} \frac{\partial \psi}{\partial x_k} x_{k+1} + \omega^{-1} \psi \right) + \sum_{i=1}^I \nu_i(u_i - \hat{u}_i), \nu_i = f_1 \tilde{v}_i. \end{aligned}$$

Далее все рассуждения проводятся аналогично. Построение регулятора для динамического объекта, модель которого задана в виде (7.9) закончена.

Замечание. Как выше было отмечено, структура (7.9) является моделью для широкого класса объектов, когда измерению доступна лишь одна наблюдаемая величина y (сама наблюдаемая величина берется в качестве x_1). Поэтому предположение о наличии тренда для переменной x_1 не является ограничительным.

Иллюстрируем применение изложенной схемы на примерах ($\omega=T$).

Пример 7.1. Пусть $\psi(x) = \sum_{k=1}^n a_k x_k, f_2(u)=u$, тогда регулятор имеет вид:

$$\begin{aligned} \dot{x}_1(t) &= x_2, \\ &\dots\dots\dots \\ \dot{x}_{n-1}(t) &= x_n, \\ \dot{x}_n(t) &= -\frac{1}{a_n} \frac{\Psi}{T} - \frac{1}{a_n} \sum_{k=1}^{n-1} a_k x_{k+1} + \sum_{i=1}^I \vartheta_i \eta_i, \\ \dot{\eta}_i &= -a_n \gamma \vartheta_i \sum_{k=1}^n a_k x_k, \quad \eta_i = (v_i - \hat{v}_i). \end{aligned}$$

Пример 7.2. Пусть дана модель ДО в виде ($n=2$):

$$\begin{aligned} \dot{x}_1(t) &= x_2, \\ \dot{x}_2(t) &= f(x, \theta, \xi) + u(t), \end{aligned} \tag{7.22}$$

и пусть найдена функциональная зависимость (гл. 2, реконструкция ВР) $x_{1tr}(t) = h(t)$. Подставим в (7.22):

$$\begin{aligned} \dot{x}_{1tr}(t) &= \dot{h}(t) = x_2, \\ \dot{x}_2(t) &= \ddot{h}(t) = -\frac{1}{a_2} \frac{\Psi}{T} - \frac{1}{a_2} a_1 x_2 + \sum_{i=1}^I \vartheta_i \eta_i = -\frac{1}{a_2} \frac{\Psi}{T} - \frac{1}{a_2} a_1 \dot{h}(t) + \sum_{i=1}^I \vartheta_i \eta_i, \end{aligned}$$

получим соотношение для весовых коэффициентов:

$$\sum_{i=1}^I v_i \eta_i = \ddot{h}(t) + \frac{1}{a_2} \frac{\Psi}{T} + \frac{1}{a_2} \dot{h}(t).$$

Таким образом, динамический регулятор имеет вид:

$$\begin{aligned} \dot{x}_1(t) &= x_2, \\ \dot{x}_2(t) &= -\frac{1}{a_2} \frac{\Psi}{T} - \frac{1}{a_2} a_1 x_2 + \sum_{i=1}^I v_i \eta_i, \\ \sum_{i=1}^I v_i \eta_i &= \ddot{h}(t) + \frac{1}{a_2} \frac{\Psi}{T} + \frac{1}{a_2} \dot{h}(t), \\ \dot{\eta}_i &= -a_2 \gamma v_i (a_1 x_1 + a_2 x_2), \quad i = \overline{1, I}. \end{aligned}$$

Рассмотрим простой пример модели, для которого в [7] был построен нечеткий регулятор с целью сравнения эффективности методик (времени формирования управляющего воздействия и длительности переходных процессов).

Пример 7.3. Пусть дана модель ДО в виде:

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= f + u, f = \theta_1 x_2 + \theta_2 \text{sign}(x_1) + \xi\end{aligned}$$

Выберем целевое многообразие в виде:

$$\psi = c_1 x_1 + c_2 x_2 - c, \quad c, c_1, c_2 = \text{const}.$$

Согласно полученному динамическому регулятору (7.21), в котором роль входных воздействий играют величины $\eta_i, i = \overline{1, I}$, исходная система для 2-х состояний ($I=2$) и значений параметров $\theta_1=-0.1, \theta_2=1$ будет иметь вид:

$$\begin{aligned}\dot{x}_1(t) &= x_2, \\ \dot{x}_2(t) &= -\frac{x_1}{T} - \left(\frac{1}{T} + 1\right)x_2 + \vartheta_1 \eta_1 + \vartheta_2 \eta_2, \\ \dot{\eta}_1 &= -\gamma \vartheta_1 c_2 (c_1 x_1 + c_2 x_2 - c), \\ \dot{\eta}_2 &= -\gamma \vartheta_2 c_2 (c_1 x_1 + c_2 x_2 - c),\end{aligned}$$

или в векторно-матричном виде:

$$\begin{aligned}\dot{x} &= A(x)x + G, \quad x = (x_1, x_2, \eta_1, \eta_2)^T, \\ A(x) &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{1}{T} & -\left(\frac{1}{T} + 1\right) & \vartheta_1 & \vartheta_2 \\ -\gamma \vartheta_1 c_1 c_2 & -\gamma \vartheta_1 c_2^2 & 0 & 0 \\ -\gamma \vartheta_2 c_1 c_2 & -\gamma \vartheta_2 c_2^2 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 \\ 0 \\ \gamma \vartheta_1 c_2 c \\ \gamma \vartheta_2 c_2 c \end{bmatrix}.\end{aligned}$$

После применения одношаговой разностной схемы Эйлера исходная система примет вид:

$$\begin{aligned}x[k+1] &= F(x[k])x[k] + H, \\ F(x) = E_2 + \tau_0 A(x[k]) &= \begin{bmatrix} 1 & \tau_0 & 0 & 0 \\ -\frac{\tau_0}{T} & 1 - \tau_0 \left(\frac{1}{T} + 1\right) & \tau_0 \vartheta_1 & \tau_0 \vartheta_2 \\ -\tau_0 \gamma \vartheta_1 c_1 c_2 & -\tau_0 \gamma \vartheta_1 c_2^2 & 1 & 0 \\ -\tau_0 \gamma \vartheta_2 c_1 c_2 & -\tau_0 \gamma \vartheta_2 c_2^2 & 0 & 1 \end{bmatrix}, \quad H = \tau_0 G = \begin{bmatrix} 0 \\ 0 \\ \tau_0 \gamma \vartheta_1 c_2 c \\ \tau_0 \gamma \vartheta_2 c_2 c \end{bmatrix},\end{aligned}$$

или по координатам:

$$x_1[k+1] = x_1[k] + \tau_0 x_2[k],$$

$$x_2[k+1] = -\frac{\tau_0}{T} x_1[k] + \left(1 - \tau_0 - \frac{\tau_0}{T}\right) x_2[k] + \tau_0 \mathcal{G}_1[k] \eta_1[k] + \tau_0 \mathcal{G}_2[k] \eta_2[k],$$

$$\eta_1[k+1] = -\tau_0 \gamma \mathcal{G}_1[k] c_2 (c_1 x_1[k] + c_2 x_2[k]) + \eta_1[k] + \tau_0 \gamma \mathcal{G}_1 c_2 c,$$

$$\eta_2[k+1] = -\tau_0 \gamma \mathcal{G}_2[k] c_2 (c_1 x_1[k] + c_2 x_2[k]) + \eta_2[k] + \tau_0 \gamma \mathcal{G}_2 c_2 c.$$

Коэффициенты \mathcal{G}_i , $i = \overline{1, I}$ выбираем из условия:

$$x_1[k+1] = x_{1tr}[k+1], \quad x_2[k+1] = x_{2tr}[k+1], \quad \lim_{k \rightarrow \infty} (x_{1tr}[k+1] + x_{2tr}[k+1]) = \psi.$$

Для рассматриваемого примера из выполнения следующего требования:

$$\begin{cases} x_{1tr}[k+1] = x_1[k] + \tau_0 x_2[k], \\ x_{2tr}[k+1] = -\frac{\tau_0}{T} x_1[k] + \left(1 - \tau_0 - \frac{\tau_0}{T}\right) x_2[k] + \tau_0 \mathcal{G}_1[k] \eta_1[k] + \tau_0 \mathcal{G}_2[k] \eta_2[k], \end{cases} \quad (7.23)$$

и в зависимости от номера измеряемой величины можно получить разные законы для определения искомым величин.

1. Пусть измеряется координата x_2 . Тогда сложив оба уравнения, предварительно умножив 1-е на $\frac{\tau_0}{T}$, получим искомые соотношения для коэффициентов

$$\mathcal{G}_1[k], \mathcal{G}_2[k]: \tau_0 \mathcal{G}_1[k] \eta_1[k] + \tau_0 \mathcal{G}_2[k] \eta_2[k] = \frac{\tau_0}{T} x_{1tr}[k+1] + x_{2tr}[k+1] - x_2[k] \left(1 - \tau_0 - \frac{\tau_0}{T} + \frac{\tau_0^2}{T}\right).$$

2. Пусть измеряется координата x_1 (что более естественно для прикладных задач). Выразив $x_2[k]$ из 1-го уравнения (7.23) $x_2[k] = \frac{1}{\tau_0} (x_{1tr}[k+1] - x_1[k])$ получим искомое соотношение для случая неизмеряемой 2-й координаты:

$$\begin{aligned} \mathcal{G}_1[k] \eta_1[k] + \mathcal{G}_2[k] \eta_2[k] &= \frac{1}{T} x_{1tr}[k+1] + \frac{1}{\tau_0} x_{2tr}[k+1] - \\ &- \frac{1}{\tau_0} (x_{1tr}[k+1] - x_1[k]) \left(\frac{1}{\tau_0} - \frac{(1 - \tau_0)}{T} - 1 \right). \end{aligned}$$

3. Для случая $I=1$ (что соответствует отсутствию априорной информации об особенностях ДО) весовой коэффициент будет однозначно определяться из

$$\text{выражения: } \mathcal{G}_1[k] = \frac{1}{\tau_0 \eta_1[k]} \left(\frac{\tau_0}{T} x_{1tr}[k+1] + x_{2tr}[k+1] - x_2[k] \left(1 - \tau_0 - \frac{\tau_0}{T} + \frac{\tau_0^2}{T} \right) \right).$$

Тогда динамический регулятор имеет вид:

$$x_1[k+1] = x_1[k] + \tau_0 x_2[k],$$

$$x_2[k+1] = -\frac{\tau_0}{T} x_1[k] + \left(1 - \tau_0 - \frac{\tau_0}{T}\right) x_2[k] + \tau_0 \mathcal{G}[k] \eta_1[k],$$

$$\mathcal{G}[k] = \frac{1}{\tau_0 \eta_1[k]} \left(\frac{\tau_0}{T} x_{1tr}[k+1] + x_{2tr}[k+1] - x_2[k] \left(1 - \tau_0 - \frac{\tau_0}{T} + \frac{\tau_0^2}{T}\right) \right),$$

$$\eta_1[k+1] = -\tau_0 \gamma \mathcal{G}[k] c_2 (c_1 x_1[k] + c_2 x_2[k]) + \eta_1[k] + \tau_0 \gamma \mathcal{G}[k] c_2 c.$$

4. Для случая задания целевых траекторий только для части координат ($x_1[k+1] = x_{1tr}[k+1]$) получаем следующий динамический регулятор:

$$x_1[k+1] = x_1[k] + \tau_0 x_2[k],$$

$$x_2[k+1] = -\frac{\tau_0}{T} x_1[k] + \left(1 - \tau_0 - \frac{\tau_0}{T}\right) x_2[k] + \tau_0 \mathcal{G}[k] \eta_1[k],$$

$$\mathcal{G}[k] = \frac{1}{\tau_0^2 \eta_1[k]} \left(x_{1tr}[k+2] - \left(x_1[k+1] - \frac{\tau_0^2}{T} x_1[k] \right) - \left(\tau_0 - \tau_0^2 - \frac{\tau_0^2}{T} \right) x_2[k] \right),$$

$$\eta_1[k+1] = -\tau_0 \gamma \mathcal{G}[k] c_2 (c_1 x_1[k] + c_2 x_2[k]) + \eta_1[k] + \tau_0 \gamma \mathcal{G}[k] c_2 c.$$

Приведем для вышеобозначенных случаев 3, 4 результаты численного моделирования при разных начальных условиях (рис. 7.3), параметрах настройки регулятора (рис. 7.4 - рис. 7.6), при разных аппроксимациях тренда (рис. 7.3 - рис. 7.6) дискретных значений фазовых траекторий (временных рядов).

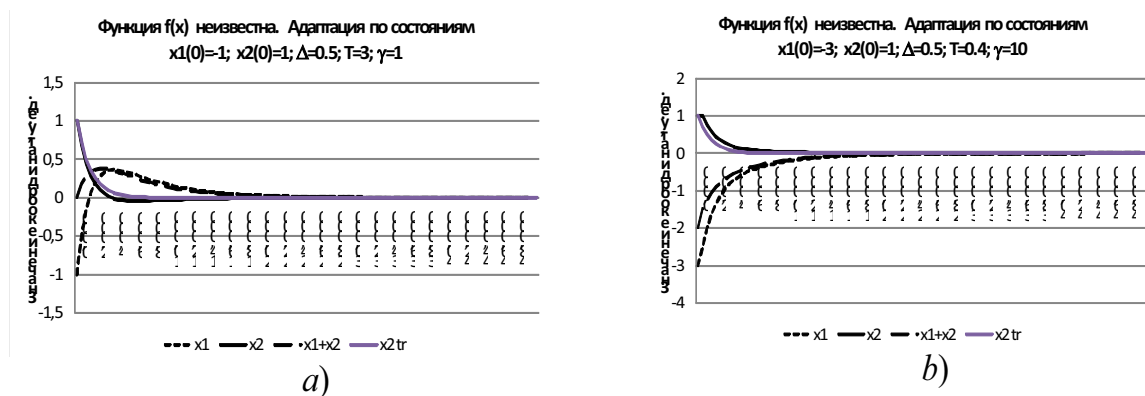


Рис 7.3. Влияние начальных условий: а) $x_1(0)=-1$; $x_2(0)=1$; б) $x_1(0)=-1.5$; $x_2(0)=1$ на вид переходного периода с функциональной зависимостью целевых траекторий $x_{1tr}=-x_{2tr}$, $x_{2tr}=\exp(-\lambda x)$, $\lambda=0.7$ и адаптацией по x_1, x_2

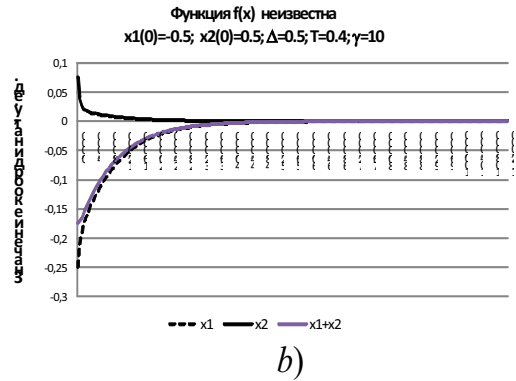
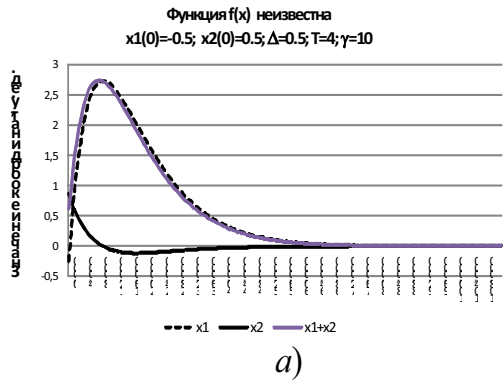


Рис. 7.4. Влияние разных параметров настройки регулятора а) $T=4$; б) $T=0.4$ на вид переходного периода с функциональной зависимостью целевых траекторий $x_{1tr} = -x_{2tr}$, $x_{2tr} = \exp(-\lambda x)$, $\lambda=0.7$ и адаптацией по x_1, x_2

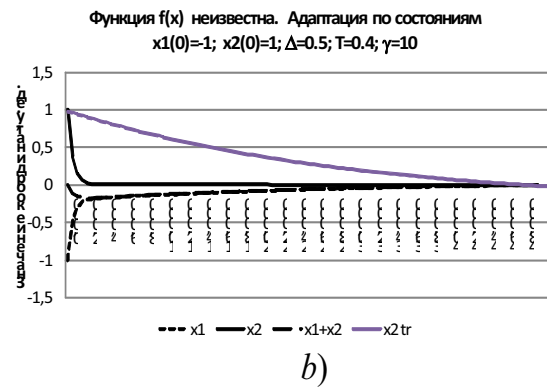
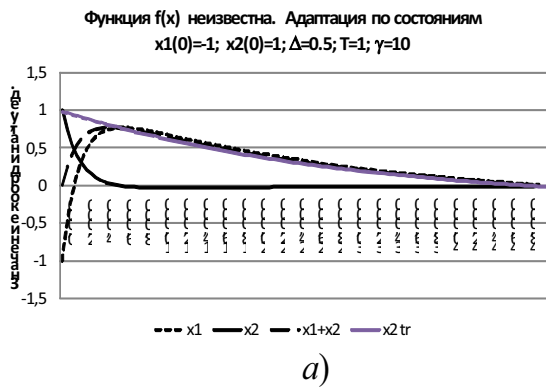


Рис. 7.5. Влияние разных параметров настройки регулятора а) $T=4$; б) $T=0.4$ на вид переходного периода с функциональной зависимостью целевых траекторий $x_{1tr} = -x_{2tr}$, $x_{2tr} = -2E-11x_5 + 2E-08x_4 - 5E-06x_3 + 0.0006x_2 - 0.0398x + 0.9812$ и адаптацией по x_1, x_2

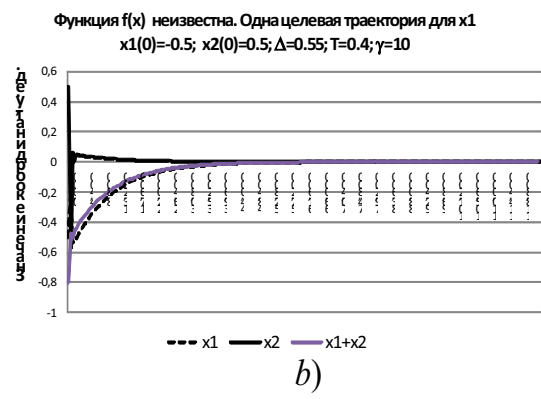
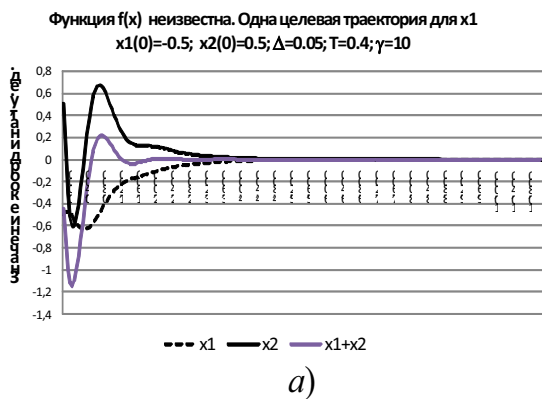


Рис. 7.6. Задание целевой траектории x_{1tr} только для одной координаты устойчиво выводит систему (7.7) на заданное многообразие: а) $\Delta=0.05$; б) $\Delta=0.55$ (Δ - интервал дискретизации)

Численное сравнение с нечетким регулятором проведено на ряде примеров с эффектом существенного экономии времени управления (десятки раз) и с при-

емлемым временным интервалом переходного периода, зависящим от параметров настройки динамического регулятора.

7.4. Регулятор для управления слабоформализованным ДО в пространстве наблюдений

Возможность приведения любой системы ОДУ к форме (7.8) доказана Такенсом [18], однако, во-первых, такое представление осуществляется, как правило, за счет увеличения размерности ($D > n$); во-вторых, не всегда можно получить функцию F в явном виде.

Рассмотрим класс ДО, имеющих описание:

$$\begin{aligned} \dot{x}_j &= f_j(x_1, \dots, x_n), j = \overline{1, n-m}, \\ \dot{x}_j &= f_j(x_1, \dots, x_n) + u_j, j = \overline{n-m+1, n}, \\ z_k &= g_k(x_1, \dots, x_n) + \zeta_k, k \in \{j_1, \dots, j_s\}, j_l \in \{1, \dots, n\}, s \leq n. \end{aligned} \quad (7.24)$$

$x(t) = (x_1, \dots, x_n) \in R^n$ - вектор переменных состояния; $\theta(t) \in R^m$ - вектор параметров; $u(t)$ - управляющий сигнал; $\xi(t), \zeta \in R$ - возмущения, $f_i(\cdot), g(\cdot)$ - неизвестные нелинейные (в общем виде) функции.

Требуется осуществить управление в пространстве состояний объекта с использованием т.н. макропеременных $\psi_l: R^n \rightarrow R, \psi_l \in C^1, l = 1, \dots, m, m < n$, равенство нулю которых задает желаемые инвариантные многообразия, а именно, сформировать управляющие воздействия, переводящие объект из заданного начального состояния x_0 в окрестность $\psi(x) = 0$ и доставляющие минимум функционалу вида (с учетом m -мерности управления):

$$\Phi = \int_0^{\infty} \sum_{i=1}^m [\phi_i^2(\psi_i) + T_i^2 \dot{\psi}_i^2(t)] dt \quad (7.25)$$

Теорема 7.2. Для класса динамических объектов, описываемых моделью (7.24) с вектором состояний $x \in R^n$, вектором управления $u \in R^m$, $\zeta \in R^p$ - вектором возмущений, $f \in R^n$ - неизвестной нелинейной (в общем виде) вектор-функции, существует закон управления в виде:

$$u(\psi) = u(x) = \mathcal{G}^T u, \quad \mathcal{G} = \|\mathcal{G}_{li}\|_{l \times 1}, \quad u = \|u_{li}\|_{l \times 1}, \quad (7.26)$$

с правилом определения управляющих воздействий u_i и весовых коэффициентов \mathcal{G}_i управляющих воздействий в i -м состоянии динамического объекта (7.24) в виде:

$$\dot{u}_i = -\gamma g_i \frac{\psi \partial \psi}{\partial x_j}, \quad \frac{\partial \psi}{\partial x_j} \neq 0, \quad i = \overline{1, I}, \quad j = \overline{1, n}, \quad \gamma = \text{const} > 0, \quad (7.27)$$

обеспечивающий

1) перевод объекта управления (7.24) из произвольного начального состояния $x_0 \in R^n$ в некоторой области фазового пространства в заданное состояние $\psi(x)=0$ и его стабилизацию в некоторой окрестности инвариантного многообразия $\psi(x)=0$ с макропеременной $\psi(x)$, ограниченной условием $\frac{\partial \psi}{\partial x_j} \neq 0$ и

удовлетворяющей соотношению

$$\psi(x) = \varphi(z, \dot{z}), \quad z = y - y_{et}, \quad \alpha > 0, \quad (7.28)$$

где y - наблюдаемое значение выходной переменной, y_{et} - заданное целевое значение (значение целевой траектории в т.ч.);

2) доставление минимума на траекториях движения замкнутой системы сопровождающему оптимизирующему функционалу вида (7.25);

3) асимптотическую устойчивость объекта (7.24).

Доказательство теоремы 7.2 аналогично доказательству теоремы 7.1, но более громоздко.

Вышеизложенные результаты (теоремы и примеры) представляют собой подход, идейно близкий подходу к формированию управления, основанного на применении самоорганизующегося оптимального регулятора с экстраполяцией (СОРЭ) А.А. Красовского, однако реализация регулятора принципиально другая и основана на распознавании состояний ДО (как множества возможных значений) и применении этого знания для управления сложным объектом с возможными параметрическими флуктуациями, присутствием немоделируемой нелинейной динамики, нестационарными ошибками измерения. В СОРЭ ввиду отсутствия достоверной математической модели ДО можно использовать его информационную модель в виде некоторых полиномов, характеризующих динамику соответствующих выходов. Величина шага поступления информации, протяженность цикла, необходимая для определения информационного полинома, и связанное с последним требованием оптимальное время экстраполяции определяются динамикой ДО.

7.5. Аналитическое конструирование нелинейных систем с наблюдателем состояния на основе развития АКАР

Наиболее сложной проблемой синтеза нелинейных систем при неполной информации о векторе состояния объекта и действующих на него возмущений является аналитическое конструирование нелинейных динамических систем с

наблюдателем. На данный момент формализованные и конструктивные общие методы синтеза наблюдателя отсутствуют как в отечественной, так и зарубежной научной литературе (см., например, обзор в [158]).

Классической постановкой задачи конструирования нелинейных динамических систем с наблюдателем в АКАР является следующая.

Задана математическая модель объекта управления в виде:

$$\begin{aligned} \dot{x} &= \mathbf{a}(x)x + \mathbf{B}(x)u + \mathbf{H}(x)w, \\ y &= \mathbf{c}(x)x, \end{aligned} \quad (7.29)$$

где x – n -вектор состояния, u – m -вектор управления, w – μ -вектор воздействия, y – ν -вектор наблюдаемых переменных.

Векторы $\mathbf{a}(x)$, $\mathbf{c}(x)$ и матрицы $\mathbf{B}(x)$ и $\mathbf{H}(x)$ предполагаются непрерывными и дифференцируемыми.

Воздействие w принадлежит классу детерминированных воздействий заданной формы и определяется однородным дифференциальным уравнением:

$$\dot{w} = \mathbf{G}w, \quad w(0) \in W, \quad (7.30)$$

где \mathbf{G} – заданная матрица, а $W \subset R^\mu$ – компактное множество.

Задача заключается в синтезе динамического регулятора, обеспечивающего асимптотическую устойчивость системы (7.29) в целом с заданным временем переходных процессов в ней и инвариантность выхода системы y к воздействию w заданной формы, определенному уравнением (7.30).

Обобщим схему построения наблюдателя по АКАР для случая, когда модель изменения неизмеряемых параметров и воздействий неизвестна.

Пусть модель ДО задана в виде:

$$\begin{aligned} \dot{x} &= f(x, \theta, u, t, \nu), \\ y &= h(x, t), \end{aligned} \quad (7.31)$$

$x \in R^n$ – вектор состояний, $\theta \in R^k$, – вектор параметров, ν – вектор параметров, подверженный неизмеряемым изменениям в процессе функционирования ДО, или возмущающим воздействиям, $u \in R^m$, $m < n$, – вектор управления, $f(\cdot)$, $h(\cdot)$ – непрерывные нелинейные вектор-функции, $y \in R^k$, – вектор выходных переменных (измерений).

Как и в АКАР задача будет заключаться в синтезе динамического регулятора

$$\begin{aligned} \dot{z} &= \mathbf{R}(y, z), \\ u &= \mathbf{u}(y, z) \end{aligned} \quad (7.32)$$

(z – вектор состояния), обеспечивающего асимптотическую устойчивость системы (7.31), (7.32) в целом с заданным временем переходных процессов в ней.

Для решения задачи используется метод инвариантных многообразий [70, 111, 149] и принцип разделения. В соответствии со вторым, на первом этапе в предположении полной наблюдаемости векторов x и v определяется управление по состоянию $u(x, v)$, обеспечивающее асимптотическую устойчивость системы (7.31) в области или в целом. На втором этапе для вектора v и ненаблюдаемых компонент вектора x строится асимптотический наблюдатель. Динамический регулятор (7.32) получается путем замены в полученном на первом этапе выражении для ненаблюдаемых переменных их асимптотическими оценками, формируемыми наблюдателем.

Пусть модель ДО (7.31) представима в виде:

$$\begin{aligned} \dot{x}_1 &= f_1(x, \theta, u), \\ &\dots\dots\dots \\ \dot{x}_{k+1} &= f_{k+1}(x, \theta, u, v), \\ &\dots\dots\dots \\ \dot{x}_{k+s} &= f_{k+s}(x, \theta, u, v), \\ &\dots\dots\dots \\ \dot{x}_n &= f_n(x, \theta, u), \end{aligned} \tag{7.33}$$

где $f_i(\cdot)$ – известные функции, $v = (v_1, \dots, v_s) \in R^s$, - вектор параметров с неизмеряемыми флуктуациями, влияющими на качество управления.

Утверждение 7.1. Если уравнения измеряемых координат, содержащие неизвестные параметры (возмущения) однозначно разрешимы относительно оцениваемых параметров (возмущений), то для расширенной системы (7.34) за счет полученных уравнений ($v_1 = x_{n+1}, \dots, v_s = x_{n+s}$):

$$\begin{aligned} \dot{x}_1 &= f_1(x, \theta, u), \\ &\dots\dots\dots \\ \dot{x}_n &= f_n(x, \theta, u), \\ \dot{x}_{n+1} &= \ddot{x}_{k+1} - \frac{d}{dt} f_{k+1}(x, \theta, u), \\ &\dots\dots\dots \\ \dot{x}_{n+s} &= \ddot{x}_{k+s} - \frac{d}{dt} f_{k+s}(x, \theta, u), \end{aligned} \tag{7.34}$$

при фиксированном управлении u и векторе x^1 измеряемых компонент вектора x переменных состояния существует асимптотический наблюдатель для оцениваемых параметров (возмущений) в виде:

$$\dot{z} = L(x^1)z - L(x^1) \int_0^{x^1} \Gamma(x^1) dx^1 - \alpha(x^1) - \beta(x^1)u, \quad \varphi(x^1, \hat{g}) = \int_0^{x^1} \Gamma(x^1) dx^1 - z, \tag{7.35}$$

где $L(x^1)$ – нижеопределенная матрица, а функции $\varphi(x^1, g)$, $\Gamma(x^1)$, $\alpha(x^1)$, $\beta(x^1)$ удовлетворяют уравнениям:

$$\begin{aligned} \frac{\partial \varphi}{\partial \mathbf{x}^1} \mathbf{B}_1(\mathbf{x}^1, \mathbf{g}) + \frac{\partial \varphi}{\partial \mathbf{g}} \mathbf{B}_2(\mathbf{x}^1, \mathbf{g}) &= \Gamma(\mathbf{x}^1) \mathbf{B}_1(\mathbf{x}^1, \mathbf{g}) + \beta(\mathbf{x}^1), \\ \frac{\partial \varphi}{\partial \mathbf{x}^1} \mathbf{A}_1(\mathbf{x}^1, \mathbf{g}) + \frac{\partial \varphi}{\partial \mathbf{g}} \mathbf{A}_2(\mathbf{x}^1, \mathbf{g}) - \mathbf{L}(\mathbf{x}^1) \varphi(\mathbf{x}^1, \mathbf{g}) &= \Gamma(\mathbf{x}^1) \mathbf{A}_1(\mathbf{x}^1, \mathbf{g}) + \alpha(\mathbf{x}^1), \end{aligned} \quad (7.36)$$

а функция $\varphi(\mathbf{x}^1, \mathbf{g})$ непрерывна и дифференцируема по своим аргументам.

Замечания.

1. Исходная неопределенность в системе (7.33) относительно неизмеряемого вектора параметров ν (и/или возмущения) в системе (7.34) переходит к другой переменной $\ddot{\mathbf{x}}$, которая будет интерпретироваться как внешнее возмущение (с меньшей неопределенностью, если сама координата x является измеряемой).

2. Переменная, в уравнение которой входит оцениваемый параметр (и/или внешнее воздействие), должна быть необходимо наблюдаемой (измеряемой).

Доказательство опирается на работы А.А. Колесникова и его учеников [70, 149] и одновременно является схемой построения наблюдателя, иллюстрация для которого приведена ниже на примере наблюдателя для одной из форм модели асинхронного двигателя (п. 7.6).

Приведем уравнение объекта (7.34) к виду, удобному для построения наблюдателя. С этой целью представим вектор состояния объекта \mathbf{x} в виде:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}^{1T} & \mathbf{x}^{2T} \end{bmatrix}, \quad (7.37)$$

где вектор \mathbf{x}^1 размерности $k \geq 1$ составлен из измеряемых компонент вектора \mathbf{x} , для которых отображение

$$\mathbf{y}: \mathbf{x}^1 \rightarrow h(\mathbf{x}^1, \mathbf{x}^2),$$

определенное уравнением выхода в (7.31) является взаимно однозначным для всех $\mathbf{x}^2 \in R^{n-k}$, \mathbf{x}^2 – вектор, состоящий из неизмеряемых компонент вектора \mathbf{x} .

Образуем вектор $\mathbf{g} \in R^{n-k+s}$:

$$\mathbf{g} = \begin{bmatrix} \mathbf{x}^{2T} & \nu^T \end{bmatrix}$$

и предположим, что уравнение

$$\mathbf{y} = h(\mathbf{x}^1, \mathbf{x}^2), \quad (7.38)$$

однозначно разрешимо в явном виде относительно \mathbf{x}^2 . Без ограничения общности будем считать, что уравнения (7.31), (7.34) с учетом введенных обозначений можно записать в следующей эквивалентной форме:

$$\begin{aligned} \dot{\mathbf{x}}^1 &= \mathbf{A}_1(\mathbf{x}^1, \mathbf{g}) + \mathbf{B}_1(\mathbf{x}^1, \mathbf{g}) \mathbf{u}, \\ \dot{\mathbf{g}} &= \mathbf{A}_2(\mathbf{x}^1, \mathbf{g}) + \mathbf{B}_2(\mathbf{x}^1, \mathbf{g}) \mathbf{u}. \end{aligned} \quad (7.39)$$

Здесь \mathbf{x}^1 – полностью измеряемый, а \mathbf{g} – неизмеряемый (неизвестный) вектор размерности $l = n - k + s$.

Замечание. Требование однозначности разрешения (7.38) относительно x^1 не является критичным. В случае неоднозначности разрешения (7.38) относительно x^1 при формировании системы (7.39) следует выбирать то решение, которое приводит к более простой структуре по переменной g .

Далее, введем в рассмотрение вектор $\psi(t) \in R^l$, определяемый формулой:

$$\psi(t) = \varphi(x^1, g) - \hat{\varphi}(t), \quad (7.40)$$

где функции φ и $\hat{\varphi}$ удовлетворяют следующим условиям:

- а) $\varphi(x^1, g)$ и $\hat{\varphi}(t)$ непрерывны и дифференцируемы по своим аргументам;
- б) решение уравнения

$$\varphi(x^1, g) = \hat{\varphi}(t), \quad (7.41)$$

относительно g существует и является единственным для всех $g \in R^s$.

Рассматривая вектор ψ как функцию времени, потребуем, чтобы он удовлетворял однородному дифференциальному уравнению:

$$\dot{\psi} = L(x^1)\psi, \quad (7.42)$$

где $L(x^1)$ – матрица $l \times l$ такова, что тривиальное решение $\psi=0$ данного уравнения асимптотически устойчиво в целом ($L(x^1)$ – числовая устойчивая матрица). С учетом выражений (7.40) и (7.41) уравнение (7.42) принимает вид:

$$\frac{\partial \varphi}{\partial x^1} A_1(x^1, g) + \frac{\partial \varphi}{\partial g} A_2(x^1, g) + \left(\frac{\partial \varphi}{\partial x^1} B_1(x^1, g) + \frac{\partial \varphi}{\partial g} B_2(x^1, g) \right) u - \dot{\hat{\varphi}} = L(x^1) (\varphi(x^1, g) - \hat{\varphi}(t)). \quad (7.43)$$

Предположим, что для заданных $A_1(x^1, g)$, $A_2(x^1, g)$, $B_1(x^1, g)$ и $B_2(x^1, g)$ найдутся такие векторы $\alpha(x^1)$, $\beta(x^1)$ и матрица $\Gamma(x^1)$, не зависящие от g , что оказываются выполненными равенства (7.36).

Тогда уравнение (7.43) с учетом первого уравнения системы (7.39) примет вид:

$$\Gamma(x^1)\dot{x}^1 - \dot{\hat{\varphi}} + \alpha(x^1) + \beta(x^1)u + L(x^1)\hat{\varphi} = 0,$$

или

$$\dot{z} = L(x^1)z - L(x^1) \int_0^{x^1} \Gamma(x^1) dx^1 - \alpha(x^1) - \beta(x^1)u, \quad (7.44)$$

где

$$z = \int_0^{x^1} \Gamma(x^1) dx^1 - \hat{\varphi}, \quad (7.45)$$

При этом, согласно уравнениям (7.40) и (7.45), вектор ψ определится из соотношения:

$$\psi = \varphi(\mathbf{x}^1, \mathbf{g}) - \int_0^{x^1} \Gamma(\mathbf{x}^1) dx^1 + z, \quad (7.46)$$

Поскольку решение $\psi = 0$ уравнения (7.42) асимптотически устойчиво в целом, то асимптотическая оценка $\hat{\mathbf{g}}$ ненаблюдаемого вектора \mathbf{g} может быть найдена как решение уравнения:

$$\varphi(\mathbf{x}^1, \hat{\mathbf{g}}) = \int_0^{x^1} \Gamma(\mathbf{x}^1) dx^1 - z, \quad (7.47)$$

которое, согласно условию (7.41), является единственным решением относительно $\hat{\mathbf{g}}$ для всех $\mathbf{x}^1 \in R^k$.

Задача синтеза наблюдателя сводится, таким образом, к нахождению функций $\varphi(\mathbf{x}^1, \mathbf{g})$, $\Gamma(\mathbf{x}^1)$, $\alpha(\mathbf{x}^1)$, $\beta(\mathbf{x}^1)$, удовлетворяющих уравнениям (7.36).

Утверждение 7.1. доказано.

7.6. Иллюстративный пример построения наблюдателя для неизмеряемых координат и возмущений

Рассмотрим одну из моделей асинхронного двигателя [69]:

$$\begin{aligned} \dot{x}_1 &= -a_1 x_1 + a_2 x_3 + u_1, \\ \dot{x}_2 &= -a_1 x_2 + a_2 x_4 + u_2, \\ \dot{x}_3 &= a_3 x_1 - a_4 x_3 - p x_4 x_5, \\ \dot{x}_4 &= a_3 x_2 - a_4 x_4 + p x_3 x_5, \\ \dot{x}_5 &= a_5 x_2 x_3 - a_5 x_1 x_4 - \frac{1}{J} M_C, \end{aligned} \quad (7.48)$$

где $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5) = (\psi_{1\alpha}, \psi_{1\beta}, \psi_{2\alpha}, \psi_{2\beta}, \omega)^T$ – вектор состояния АД; $\psi_{1\alpha}, \psi_{1\beta}, \psi_{2\alpha}, \psi_{2\beta}$ – составляющие векторов потокосцепления статора и ротора; $\mathbf{u} = [u_{1\alpha}, u_{1\beta}]^T$ – вектор входных (управляющих) воздействий; $u_{1\alpha}, u_{1\beta}$ – составляющие вектора напряжения статора; коэффициенты a_i выражены через физические параметры АД (п. 9.1) и имеют вид:

$$a_1 = \frac{L_r R_s}{L_s L_r - L_m^2}; a_2 = \frac{L_m R_s}{L_s L_r - L_m^2}; a_3 = \frac{L_m L_r}{L_s L_r - L_m^2}; a_4 = \frac{R_r L_s}{L_s L_r - L_m^2}; a_5 = \frac{p L_m}{J(L_s L_r - L_m^2)}.$$

Выразим из 5-го уравнения оцениваемый параметр $M_C = a_5 J x_2 x_3 - a_5 J x_1 x_4 - J \dot{x}_5$, и, изменив порядок записи уравнений, дополним систему (7.48) шестым уравнением, обозначив $x_6 = M_C$, выполняя необходимые преобразования:

$$\begin{aligned}
\dot{x}_6 &= a_5 J \dot{x}_2 x_3 + a_5 J x_2 \dot{x}_3 - a_5 J (\dot{x}_1 x_4 + x_1 \dot{x}_4) - J \ddot{x}_5 = a_5 J \dot{x}_2 x_3 + \\
&+ a_5 J x_2 (a_3 x_1 - a_4 x_3 - p x_4 x_5) - a_5 J \dot{x}_1 x_4 - a_5 J x_1 (a_3 x_2 - a_4 x_4 + p x_3 x_5) - J \ddot{x}_5 = \\
&= a_5 J \dot{x}_2 x_3 - a_5 J x_2 (a_4 x_3 + p x_4 x_5) - a_5 J \dot{x}_1 x_4 + a_5 J x_1 (a_4 x_4 - p x_3 x_5) - J \ddot{x}_5.
\end{aligned}$$

Система (7.48) примет вид:

$$\begin{aligned}
\dot{x}_1 &= -a_1 x_1 + a_2 x_3 + u_1, & \dot{x}_1 &= -a_1 x_1 + a_2 x_3 + u_1, \\
\dot{x}_2 &= -a_1 x_2 + a_2 x_4 + u_2, & \dot{x}_2 &= -a_1 x_2 + a_2 x_4 + u_2, \\
\dot{x}_5 &= a_5 x_2 x_3 - a_5 x_1 x_4 - \frac{1}{J} x_6, & \dot{x}_5 &= a_5 x_2 x_3 - a_5 x_1 x_4 - \frac{1}{J} x_6, \\
\dot{x}_3 &= a_3 x_1 - a_4 x_3 - p x_4 x_5, & \text{или} & \dot{x}_3 = a_3 x_1 - a_4 x_3 - p x_4 x_5, & (7.49) \\
\dot{x}_4 &= a_3 x_2 - a_4 x_4 + p x_3 x_5, & \dot{x}_4 &= a_3 x_2 - a_4 x_4 + p x_3 x_5, \\
\dot{x}_6 &= a_5 J \dot{x}_2 x_3 + a_5 J x_2 \dot{x}_3 - & \dot{x}_6 &= a_5 J \dot{x}_2 x_3 - a_5 J x_2 (a_4 x_3 + p x_4 x_5) - a_5 J \dot{x}_1 x_4 + \\
&- a_5 J (\dot{x}_1 x_4 + x_1 \dot{x}_4) - J \ddot{x}_5, & & + a_5 J x_1 (a_4 x_4 - p x_3 x_5) - J \ddot{x}_5.
\end{aligned}$$

Построение наблюдателя для переменных (и идентификатора для параметров) проведем по схеме, согласно [149].

Шаг 1. Введем совокупность макропеременных (по числу идентифицируемых переменных и параметров x_3, x_4, x_6):

$$\begin{aligned}
\psi_1(t) &= \phi_1 - \hat{\phi}_1(t), \\
\psi_2(t) &= \phi_2 - \hat{\phi}_2(t), \\
\psi_3(t) &= \phi_3 - \hat{\phi}_3(t),
\end{aligned} \tag{7.50}$$

удовлетворяющих решению однородных ДУ:

$$\begin{aligned}
\dot{\psi}_1(t) - L_1 \psi_1 &= 0, \\
\dot{\psi}_2(t) - L_2 \psi_2 &= 0, \\
\dot{\psi}_3(t) - L_3 \psi_3 &= 0.
\end{aligned} \tag{7.51}$$

соответственно. Здесь функции ϕ_i зависят от наблюдаемых x_3, x_4, x_6 и измеряемых переменных (x_1, x_2, x_5) , функция $\hat{\phi}_i = \hat{\phi}_i(t)$ - (асимптотическая) оценка переменной ϕ_i , подлежащая нахождению. Обозначим $N^{ob} = \{1, 2, 5\}$.

В (7.51) подставим выражения (7.50):

$$\begin{aligned}
\sum_{j=1}^6 \frac{\partial \phi_1}{\partial x_j} \dot{x}_j - \frac{d\hat{\phi}_1}{dt} - L_1 \phi_1 + L_1 \hat{\phi}_1 &= 0, \\
\sum_{j=1}^6 \frac{\partial \phi_2}{\partial x_j} \dot{x}_j - \frac{d\hat{\phi}_2}{dt} - L_2 \phi_2 + L_2 \hat{\phi}_2 &= 0, \\
\sum_{j=1}^6 \frac{\partial \phi_3}{\partial x_j} \dot{x}_j - \frac{d\hat{\phi}_3}{dt} - L_3 \phi_3 + L_3 \hat{\phi}_3 &= 0.
\end{aligned} \tag{7.52}$$

Шаг 2. С целью исключения наблюдаемых (неизмеряемых переменных, наблюдатель для которых строится) переменных из ДУ (7.49) вводятся функции γ_i , $\Gamma_{i,j}$, $i=1, \dots, 4$, $j=1, 2, 3$, (не зависящие от неизмеряемых переменных), обеспечивающие выполнение равенств:

$$\begin{aligned}
\sum_{j=1}^6 \frac{\partial \phi_1}{\partial x_j} \dot{x}_j - L_1 \phi_1 - \Gamma_{1,1} \dot{x}_1(t) - \Gamma_{1,2} \dot{x}_2(t) - \Gamma_{1,3} \dot{x}_5(t) - \gamma_1 &= 0, \\
\sum_{j=1}^6 \frac{\partial \phi_2}{\partial x_j} \dot{x}_j - L_2 \phi_2 - \Gamma_{1,1} \dot{x}_1(t) - \Gamma_{1,2} \dot{x}_2(t) - \Gamma_{1,3} \dot{x}_5(t) - \gamma_2 &= 0, \\
\sum_{j=1}^6 \frac{\partial \phi_3}{\partial x_j} \dot{x}_j - L_3 \phi_3 - \Gamma_{1,1} \dot{x}_1(t) - \Gamma_{1,2} \dot{x}_2(t) - \Gamma_{1,3} \dot{x}_5(t) - \gamma_3 &= 0.
\end{aligned} \tag{7.53}$$

Вспомогательные промежуточные функции ϕ_i выберем в виде:

$$\phi_1 = x_1 + x_2 + x_5 + x_3, \quad \phi_2 = x_1 + x_2 + x_5 + x_4, \quad \phi_3 = x_1 + x_2 + x_5 + x_6 \tag{7.54}$$

Перепишем (7.53) с учетом представления (7.54) и системы (7.49):

$$\begin{aligned}
\sum_{k \in N^{ob}} \dot{x}_k + \dot{x}_3(t) - L_1 \left(\sum_{k \in N^{ob}} x_k + x_3 \right) - \Gamma_{1,1} \dot{x}_1(t) - \Gamma_{1,2} \dot{x}_2(t) - \Gamma_{1,3} \dot{x}_5(t) - \gamma_1 &= 0, \\
\sum_{k \in N^{ob}} \dot{x}_k + \dot{x}_4(t) - L_1 \left(\sum_{k \in N^{ob}} x_k + x_4 \right) - \Gamma_{2,1} \dot{x}_1(t) - \Gamma_{2,2} \dot{x}_2(t) - \Gamma_{2,3} \dot{x}_5(t) - \gamma_2 &= 0, \\
\sum_{k \in N^{ob}} \dot{x}_k + \dot{x}_6(t) - L_1 \left(\sum_{k \in N^{ob}} x_k + x_6 \right) - \Gamma_{3,1} \dot{x}_1(t) - \Gamma_{3,2} \dot{x}_2(t) - \Gamma_{3,3} \dot{x}_5(t) - \gamma_3 &= 0.
\end{aligned}$$

Подставим значения \dot{x}_i , $i = \overline{1,6}$ из (7.49) в данные соотношения, получим:

$$\begin{aligned}
&-a_1 x_1 + a_2 x_3 + u_1 - a_1 x_2 + a_2 x_4 + u_2 + a_5 x_2 x_3 - a_5 x_1 x_4 - \frac{1}{J} x_6 + \\
&a_3 x_1 - a_4 x_3 - p x_4 x_5 - L_1 (x_1 + x_2 + x_5 + x_3) - \\
&-\Gamma_{1,1} (-a_1 x_1 + a_2 x_3 + u_1) - \Gamma_{1,2} (-a_1 x_2 + a_2 x_4 + u_2) - \\
&-\Gamma_{1,3} \left(a_5 x_2 x_3 - a_5 x_1 x_4 - \frac{1}{J} x_6 \right) - \gamma_1 = 0,
\end{aligned}$$

$$\begin{aligned}
& -a_1x_1 + a_2x_3 + u_1 - a_1x_2 + a_2x_4 + u_2 + a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6 + \\
& + a_3x_2 - a_4x_4 + px_3x_5 - L_1(x_1 + x_2 + x_5 + x_4) - \\
& -\Gamma_{2,1}(-a_1x_1 + a_2x_3 + u_1) - \Gamma_{2,2}(-a_1x_2 + a_2x_4 + u_2) - \\
& -\Gamma_{2,3}\left(a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6\right) - \gamma_2 = 0, \\
& -a_1x_1 + a_2x_3 + u_1 - a_1x_2 + a_2x_4 + u_2 + a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6 + \\
& a_5J\dot{x}_2x_3 - a_5Jx_2(a_4x_3 + px_4x_5) - a_5J\dot{x}_1x_4 + a_5Jx_1(a_4x_4 - px_3x_5) - J\ddot{x}_5 - \\
& -L_1(x_1 + x_2 + x_5 + x_4) - \Gamma_{3,1}(-a_1x_1 + a_2x_3 + u_1) - \Gamma_{3,2}(-a_1x_2 + a_2x_4 + u_2) - \\
& -\Gamma_{3,3}\left(a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6\right) - \gamma_3 = 0.
\end{aligned} \tag{7.55}$$

Функции γ_i , $\Gamma_{i,j}$, $i=1,2$, $j=1, 2, 3$, будем выбирать такие, чтобы исключить переменные, для которых строится наблюдатель. Один из путей: исключить из ДУ все слагаемые, содержащие управления, за счет выбора функций γ_i в виде:

$$\begin{aligned}
\gamma_1 &= -\Gamma_{1,2}u_2 - \Gamma_{1,1}u_1 + u_1 + u_2, \\
\gamma_2 &= -\Gamma_{2,1}u_2 - \Gamma_{2,2}u_1 + u_1 + u_2, \\
\gamma_3 &= -\Gamma_{3,1}u_2 - \Gamma_{3,2}u_1 + u_1 + u_2.
\end{aligned}$$

Подставим найденные выражения для функций γ_i в ДУ (7.55):

$$\begin{aligned}
& -a_1x_1 + a_2x_3 - a_1x_2 + a_2x_4 + a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6 + \\
& a_3x_1 - a_4x_3 - px_4x_5 - L_1(x_1 + x_2 + x_5 + x_3) - \\
& -\Gamma_{1,1}(-a_1x_1 + a_2x_3) - \Gamma_{1,2}(-a_1x_2 + a_2x_4) - \\
& -\Gamma_{1,3}\left(a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6\right) = 0, \\
& -a_1x_1 + a_2x_3 - a_1x_2 + a_2x_4 + a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6 + \\
& + a_3x_2 - a_4x_4 + px_3x_5 - L_1(x_1 + x_2 + x_5 + x_4) - \\
& -\Gamma_{2,1}(-a_1x_1 + a_2x_3) - \Gamma_{2,2}(-a_1x_2 + a_2x_4) - \\
& -\Gamma_{2,3}\left(a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6\right) = 0,
\end{aligned} \tag{7.56}$$

$$\begin{aligned}
& -a_1x_1 + a_2x_3 - a_1x_2 + a_2x_4 + a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6 + \\
& a_5J\dot{x}_2x_3 - a_5Jx_2(a_4x_3 + px_4x_5) - a_5J\dot{x}_1x_4 + a_5Jx_1(a_4x_4 - px_3x_5) - J\ddot{x}_5 - \\
& -L_1(x_1 + x_2 + x_5 + x_4) - \Gamma_{3,1}(-a_1x_1 + a_2x_3) - \Gamma_{3,2}(-a_1x_2 + a_2x_4) - \\
& -\Gamma_{3,3}\left(a_5x_2x_3 - a_5x_1x_4 - \frac{1}{J}x_6\right) = 0.
\end{aligned}$$

Для определения функций $\Gamma_{1,j}, j=1, 2, 3$, в 1-м уравнении (7.56) приравняем к нулю выражения перед неизмеряемыми координатами и свободными членами:

$$x_3 : a_2 + a_5x_2 - a_4 - L_1 - \Gamma_{1,1}a_2 - \Gamma_{1,3}a_5x_2 = 0,$$

$$x_4 : a_2 - a_5x_1 - px_5 - \Gamma_{1,2}a_2 + \Gamma_{1,3}a_5x_1 = 0,$$

$$x_6 : \frac{1}{J} + \frac{1}{J}\Gamma_{1,3} = 0,$$

$$x_3^0, x_4^0, x_6^0 : -a_1x_1 - a_1x_2 + a_3x_1 - L_1(x_1 + x_2 + x_5) + \Gamma_{1,1}a_1x_1 + \Gamma_{1,2}a_1x_2 = 0.$$

Решая данную систему относительно $\Gamma_{1,j}, j=1, 2, 3$, получаем:

$$\Gamma_{1,3} = 1, \quad \Gamma_{1,1} = \frac{a_2 - a_4 - L_1}{a_2},$$

$$\Gamma_{1,2} = \frac{1}{(a_2 - a_1x_2)} \left(\frac{x_1}{a_2} (a_2a_3 - L_1(a_1 + a_2) - a_1a_4) - x_2(a_1 + L_1) - x_5(p + L_1) + a_2 \right).$$

Аналогично, для определения функций $\Gamma_{2,j}, j=1, 2, 3$, во 2-м уравнении (7.56) приравняем к нулю выражения перед неизмеряемыми координатами и свободными членами:

$$x_3 : a_2 + a_5x_2 + px_5 - \Gamma_{2,1}a_2 - \Gamma_{2,3}a_5x_2 = 0,$$

$$x_4 : a_2 - a_5x_1 - a_4 - L_1 - \Gamma_{2,2}a_2 + \Gamma_{2,3}a_5x_1 = 0,$$

$$x_6 : -\frac{1}{J} + \Gamma_{2,3}\frac{1}{J} = 0,$$

$$x_3^0, x_4^0, x_6^0 : -a_1x_1 - a_1x_2 + a_3x_2 - L_1(x_1 + x_2 + x_5) + \Gamma_{2,1}a_1x_1 + \Gamma_{2,2}a_1x_2 = 0.$$

Решая систему относительно $\Gamma_{2,j}, j=1, 2, 3$, получаем:

$$\Gamma_{2,3} = 1, \quad \Gamma_{2,1} = 1 + \frac{p}{a_2}x_5,$$

$$\Gamma_{2,2} = \frac{1}{(a_2 - a_1x_2)} \left(x_2(a_3 - a_1) - L_1(x_1 + x_2 + x_5 + 1) + \frac{p}{a_2}x_5a_1x_1 + a_2 - a_4 \right).$$

Аналогично, для определения функций $\Gamma_{3,j}$, $j=1, 2, 3$, в 3-м уравнении (7.56) приравняем к нулю выражения перед неизмеряемыми координатами и свободными членами:

$$x_3 : a_2 + a_5 x_2 + a_5 J \dot{x}_2 - a_4 x_2 - a_5 J p x_1 x_5 - \Gamma_{3,1} a_2 - \Gamma_{3,3} a_5 x_2 = 0,$$

$$x_4 : a_2 - a_5 x_1 - a_5 J \dot{x}_1 - a_5 J p x_2 x_5 + a_5 J a_4 x_1 - L_1 - \Gamma_{3,2} a_2 + \Gamma_{3,3} a_5 x_1 = 0,$$

$$x_6 : -\frac{1}{J} + \Gamma_{3,3} \frac{1}{J} = 0,$$

$$x_3^0, x_4^0, x_6^0 : -a_1 x_1 - a_1 x_2 - J \ddot{x}_5 - L_1 (x_1 + x_2 + x_5) + \Gamma_{3,1} a_1 x_1 + \Gamma_{3,2} a_1 x_2 = 0.$$

Решая систему относительно $\Gamma_{3,j}$, $j=1, 2, 3$, получаем:

$$\Gamma_{3,3} = 1, \quad \Gamma_{3,1} = 1 + \frac{a_5 J}{a_2} \dot{x}_2 - \frac{a_4}{a_2} x_2 - \frac{a_5 J p}{a_2} x_1 x_5,$$

$$\Gamma_{3,2} = 1 + \frac{1}{(a_2 - a_1 x_2)} \left(-a_5 J \dot{x}_1 - a_5 J p x_2 x_5 + a_5 J a_4 x_1 - J \ddot{x}_5 - L_1 (x_1 + x_2 + x_5 + 1) \right) +$$

$$+ \frac{1}{(a_2 - a_1 x_2)} \left(\frac{a_5 J}{a_2} \dot{x}_2 - \frac{a_4}{a_2} x_2 - \frac{a_5 J p}{a_2} x_1 x_5 \right) a_1 x_1.$$

Тогда уравнения (7.55) с учетом найденных коэффициентов γ_i , $\Gamma_{i,j}$, $i,j=1,2,3$, могут быть переписаны в виде:

$$\Gamma_{1,1} \dot{x}_1(t) + \Gamma_{1,2} \dot{x}_2(t) + \Gamma_{1,3} \dot{x}_5(t) - \dot{\hat{\phi}}_1(t) + \gamma_1 + L_1 \hat{\phi}_1(t) = 0,$$

$$\Gamma_{2,1} \dot{x}_1(t) + \Gamma_{2,2} \dot{x}_2(t) + \Gamma_{2,3} \dot{x}_5(t) - \dot{\hat{\phi}}_2(t) + \gamma_2 + L_2 \hat{\phi}_2(t) = 0, \quad (7.57)$$

$$\Gamma_{3,1} \dot{x}_1(t) + \Gamma_{3,2} \dot{x}_2(t) + \Gamma_{3,3} \dot{x}_5(t) - \dot{\hat{\phi}}_3(t) + \gamma_3 + L_3 \hat{\phi}_3(t) = 0.$$

Шаг 3. Осуществляем дополнительные построения наблюдателей производных для измеряемых переменных (здесь x_1, x_2, x_5): $\dot{\hat{x}}_1, \dot{\hat{x}}_2, \dot{\hat{x}}_5, \ddot{\hat{x}}_1, \ddot{\hat{x}}_2, \ddot{\hat{x}}_5$ используя адаптивную скользящую аппроксимацию на основе алгоритма 2.8.

Шаг 4. Уравнения (7.57) относительно искомым функций $\hat{\phi}_i(t)$, $i=1,2,3$ приобретут вид:

$$\Gamma_{1,1} \dot{\hat{x}}_1(t) + \Gamma_{1,2} \dot{\hat{x}}_2(t) + \Gamma_{1,3} \dot{\hat{x}}_5(t) - \dot{\hat{\phi}}_1(t) + \gamma_1 + L_1 \hat{\phi}_1(t) = 0,$$

$$\Gamma_{2,1} \dot{\hat{x}}_1(t) + \Gamma_{2,2} \dot{\hat{x}}_2(t) + \Gamma_{2,3} \dot{\hat{x}}_5(t) - \dot{\hat{\phi}}_2(t) + \gamma_2 + L_2 \hat{\phi}_2(t) = 0, \quad (7.58)$$

$$\Gamma_{3,1} \dot{\hat{x}}_1(t) + \Gamma_{3,2} \dot{\hat{x}}_2(t) + \Gamma_{3,3} \dot{\hat{x}}_5(t) - \dot{\hat{\phi}}_3(t) + \gamma_3 + L_3 \hat{\phi}_3(t) = 0.$$

Полученные выражения для $\hat{\phi}_i(t)$, $i=1,2,3$ будут представлять собой асимптотические оценки переменных $\hat{\phi}_i(t)$, выбранных в виде (7.54). Тогда искомыми оценками для переменных x_3, x_4, x_6 будут иметь вид:

$$x_3 = \hat{\phi}_1(t) - x_1 - x_2 - x_5, \quad x_4 = \hat{\phi}_2(t) - x_1 - x_2 - x_5, \quad x_6 = \hat{\phi}_3(t) - x_1 - x_2 - x_5.$$

Таким образом, в отличие от техники построения наблюдателя на основе АКАР [70], предполагающей априорное задание модели для идентификации неизмеряемых воздействий и параметрических флуктуаций, данный подход использует для представления неизвестных параметров (внешних воздействий) как дополнительных переменных исходной модели (в соответствии с принципом расширения-сжатия фазового пространства [111]) в виде эволюционных уравнений не из априорных предположений, а на основе модельных уравнений исходного ДО, и дальнейшего применения техники конструирования наблюдателя по АКАР.

Замечание. При получении решения системы (7.56) может возникнуть ситуация деления на ноль, которую можно обойти регуляризацией введением малого параметра, подобного параметру регуляризации А.Н. Тихонова³⁰, и исследовать его влияние на характер решения.

Выводы. В разделе приведена схема построения регулятора на основе выявления закономерностей в пространстве состояний и принципов регулирования и использованием этой информации. Регулятор учитывает влияние признаков, способствующих или препятствующих переходу объекта в определенное состояние, и на этой основе вырабатывающий решение об управляющем воздействии. Блок распознавания состояний и признаков, способствующих их наступлению, участвует, во-первых, в идентификации состояния объекта управления; во-вторых, в выработке управляющих воздействий.

Приведена схема построения регулятора на основе выявления закономерностей в пространстве состояний с целью применения этой информации для коррекции управления по методу АКАР. Для данного метода построен наблюдатель для неизмеряемых координат и параметров, флуктуации которых являются неизвестными гладкими функциями.

Результаты, представленные в 7-й главе, опубликованы в работах [71, 75, 144].

³⁰ Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. - М.: Наука, 1986. - 287 с.

*«Ключ ясных формул к жизни не подобран,
Но как ни бейтесь с формулами вы,
Мир будет завтра праздничным и добрым
И совершенным с ног до головы»
/Павел Антокольский/*

Глава 8. МОДЕЛИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ДЛЯ РАСПОЗНАВАНИЯ СОСТОЯНИЙ ДИНАМИЧЕСКОГО ОБЪЕКТА

Модели выявления закономерностей и реализующие их алгоритмы составляют суть практических подходов к распознаванию состояний ДО в двух интеллектуальных распознающих системах, структуры которых имеют принципиальное различие, продиктованное целями их использования.

8.1. Анализ современного состояния интеллектуальных систем для распознавания состояний динамических объектов

Среди систем для удаленного тестирования алгоритмов в различных их комбинациях и на различных данных следует отметить следующие: WEKA (<http://www.cs.waikato.ac.nz/ml/weka>), RapidMiner (<http://rapid-i.com/content/view/124/1/>), находящийся в разработке Полигон алгоритмов (<http://poligon.machinelearning.ru>), ЭЙДОС (<http://www.cs.tut.fi/~lasip/>), ДЕЛЬТА (<http://lc.kubagro.ru/aidos/>) и др. Кратко охарактеризуем основные из существующих на данный момент.

Адаптивная система анализа и прогнозирования состояний сложных систем «ДЕЛЬТА» (<http://lc.kubagro.ru/aidos/>) как инструмент реализации принципов системного анализа в АСУ включает реализацию следующих этапов: разработка информационной модели объекта управления; решение задач идентификации состояния объекта управления и выработки управляющего воздействия; анализ информационной модели предметной области. Однако, в данной системе не решены вопросы обеспечения взаимодействия системы с Windows-приложениями; адаптация модели к изменениям предметной области; вопросы интерпретации результатов и использование внешних моделей, в т.ч. экспертных; не оценена надежность принимаемых решений и, что существенно, используемая методология не соответствует скоростным режимам работы технических систем.

Проект LASIP (Local Approximations in Signal and Image Processing, <http://www.cs.tut.fi/~lasip/>), реализованный в программной среде MATLAB, посвящен исследованию широкого класса новых эффективных адаптивных методов обработки сигналов. В проекте реализованы новые статистические методы восстановления зашумленного и размытого одномерного сигнала, графических

изображений, видео и 3D микроскопий. LASIP относится к загружаемым (настольным) приложениям, зависящим от сторонних пакетов прикладных программ.

К настольным приложениям относятся и WAVELAB 850 (набор Matlab функции, которые были использованы авторами в различных алгоритмах вейвлет-анализа <http://www-stat.stanford.edu/~wavelab/>); BEAMLAB 200 (включает в себя около 900 Matlab файлов, данных, а также демонстрационные скрипты, <http://www-stat.stanford.edu/~beamlab/>); SYMMLAB 090 (около 100 Matlab файлов <http://www-stat.stanford.edu/~symmlab/>).

К классу независимых приложений относятся: Advanced Signal Processing Toolkit Demo (<http://zone.ni.com>); Mealy and Moore Sequence Recognition Demo, (<http://www.mathworks.com>); Computer Science Lab of Université Paris 6 (<http://www-connex.lip6.fr/~artieres>); Pattern Recognition Demo (http://cgi.di.uoa.gr/~stpatrec/Demo_Download.htm).

Одним из фундаментальных средств анализа информации являются логические тесты [56]. Основателем успешно развиваемого направления, связанного с их применением в рамках тестового подхода к распознаванию образов, является Ю.И. Журавлев.

Интеллектуальные системы на базе методов распознавания образов. Система ЛОРЕГ³⁰ предназначена для решения задач распознавания образов и выявления логических закономерностей в знаниях, а также голосования по ним; экспертная система логического распознавания ЭКСИЛОР³¹ работает с разнообразными объектами, описываемыми точками в пространстве многозначных переменных, и представлением знаний как системы закономерностей в форме секционированной булевой матрицы дизъюнктов.

К инструментальным средствам разработки интеллектуальных систем, реализующих математические модели и методы интеллектуального анализа информации относятся следующие:

AgSDK и MAS DK – для разработки мультиагентных систем;

АТ-ТЕХНОЛОГИЯ – для поддержки разработки веб-ориентированных интегрированных экспертных систем;

SETRIAN – для конструирования динамических систем поддержки принятия решений;

ПРИЗМА – для создания сред анализа и распознавания изображений и ряд других;

³⁰ Богомолов В.П., Виноградов А.П., Борончихин В.А., Журавлев Ю.И. и соавт. Программная система распознавания ЛОРЕГ. – Москва: Изд-во ВЦ РАН, 1998. – 64 с.

³¹ Закревский А.Д. ЭКСИЛОР – экспертная система логического распознавания// Управляющие системы и машины. – 1992. – № 5/6. – С. 118-125.

ИИС (интеллектуальное инструментальное средство) ИМСЛОГ – интегрированная среда разработки и отладки интеллектуальных тестовых распознающих систем.

8.2. Концептуальная схема и структура инструментального средства ИИС ИМСЛОГ

Архитектура ИИС ИМСЛОГ. Система [24] ориентирована на семейство операционных систем Windows. Его архитектура представляет собой иерархическую структуру программных модулей. Один модуль является резидентным и называется ядром. Ядро имеет собственную систему команд, выполняет функции координирующего центра и обеспечивает: автоматическое подключение и регистрацию всех других программных модулей; динамическую загрузку модулей; передачу управления от одного модуля к другому; информационную связь между модулями; контроль корректности выполнения программных модулей.

Все остальные программные модули являются динамически подключаемыми (плагины). По назначению плагины подразделяются на функциональные модули, модули системных данных и модуль интеллектуального интерфейса с пользователем (рис. 8.1). Функциональные модули реализуют отдельные подсистемы и функции создаваемых интеллектуальных систем, включая различные подходы и алгоритмы тестового распознавания образов. Модули системных данных хранят различные информационные структуры, используемые при работе. Модуль интеллектуального интерфейса обеспечивает интерактивное взаимодействие с пользователем на ограниченном естественном языке.

Основные связи между компонентами ИИС ИМСЛОГ представлены на рис. 8.1.

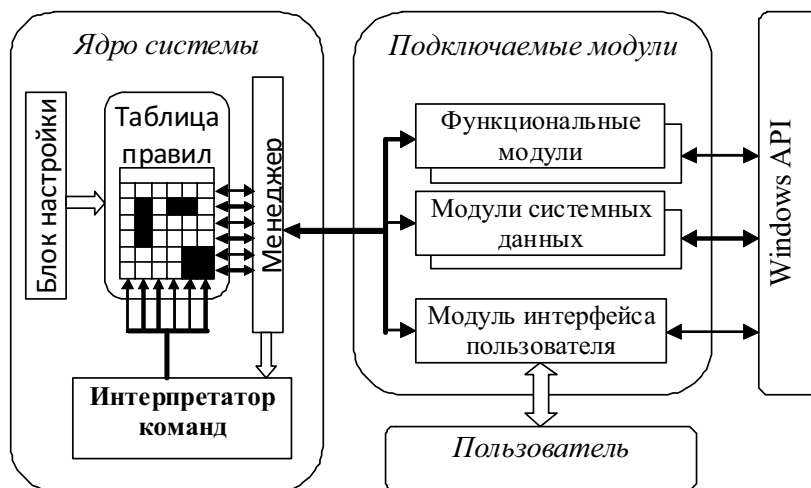


Рис. 8.1. Структурная схема ИИС ИМСЛОГ

Ядро состоит из блока конфигурации, интерпретатора системных команд, таблицы правил и менеджера подключаемых модулей.

Блок конфигурации обеспечивает хранение информации о текущей конфигурации интеллектуальной системы, о наличии подключаемых модулей, о реализованном в данной конфигурации функциональном наборе и о ресурсах, используемых плагинами в процессе работы. Ядро использует блок конфигурации при начальной загрузке, инициализации данных ядра, в процессе подключения плагинов, при чтении либо сохранении изменений в текущей конфигурации, а также при доступе к используемым ресурсам. Конфигурация сохраняется в отдельном файле и используется блоком конфигурации на протяжении всего сеанса работы.

Встроенный интерпретатор системных команд предназначен для корректного анализа поступающих в ядро команд, их сортировки и выработки управляющего воздействия для передачи управления модулю, выполняющему функции, сопоставленные данной команде. Если несколько модулей зарегистрированы для выполнения одной системной команды, то интерпретатор команд выбирает наиболее подходящую последовательность этапов ее выполнения и передает управление одному из модулей-претендентов на ее выполнение. При поступлении незнакомой команды, интерпретатор игнорирует ее и заносит соответствующее сообщение в системный протокол работы.

Таблица правил создается на основе блока конфигурации на начальном этапе работы и может изменяться на основе текущего состояния системы в процессе работы.

Менеджер подключаемых модулей является основной действующей компонентой ядра. Он обеспечивает загрузку и подключение динамических загружаемых программных модулей, контроль за модулями в процессе их работы либо в режиме отладки и передачу управления модулю, который реализует команду, поступившую в интерпретатор команд. При этом информацию о наличии подключаемого модуля и соответствии его определенной команде, а также необходимость инициализации соответствующих системных переменных и регистров флагов менеджер получает, соответственно, из блока конфигурации и таблицы правил.

Все плагины содержат специальный блок регистрации в ядре, необходимый для автоматического подключения к создаваемой прикладной интеллектуальной системе. Если плагину в процессе работы требуется устанавливать межмодульные связи, то он использует команды ядра. По способу вызова менеджером модулей плагины подразделяются на следующие типы:

- 1) модули, работающие автономно;
- 2) звено в цепочке выполняемых модулей;
- 3) модули, работающие по принципу «один из множества».

Программный код, содержащийся в теле модуля первого типа, предназначен для безусловного выполнения системной команды независимо от ситуации. При регистрации нескольких плагинов (рис. 8.2) данного типа на одну системную команду управление во время работы передается тому модулю, который зарегистрирован в блоке конфигурации и подключается посредством блока настройки при запуске системы. Другие плагины первого типа могут быть запущены только путем явного запроса к менеджеру модулей.

Результат выполнения модулей второго типа зависит от состояния системы и может быть различным. Модули данного типа используются, если требуется, чтобы какая-либо команда выполнялась несколько раз различными способами. В этом случае выполнение соответствующей команды производится рекурсивно до получения кода возврата, означающего завершение цепочки, либо до завершения работы последнего подключаемого модуля, сопоставленного этой команде. При наличии нескольких модулей третьего типа, зарегистрированных на одну системную команду, производится выбор одного из них и передача управления выбранному модулю. Выбор осуществляется путем выполнения ряда шагов. Переход к следующему шагу осуществляется в том случае, если предыдущий не привел к однозначному определению последовательности этапов выполнения команды. В первую очередь производится анализ блока конфигурации. Если в нем явным образом не осуществлена привязка одного из модулей к команде, то для выбора модуля используется таблица правил. Если с ее помощью невозможен однозначный выбор модуля, то ядро генерирует системную команду выбора модуля, выполняемую с помощью специального плагина. Данный плагин создается непосредственно при проектировании прикладной интеллектуальной системы, и критерии выбора зависят от решаемой задачи. В случае неуспешного выполнения этого плагина либо при его отсутствии, система создает специальное окно диалога и предоставляет возможность дальнейшего управления ходом выполнения команды пользователю.

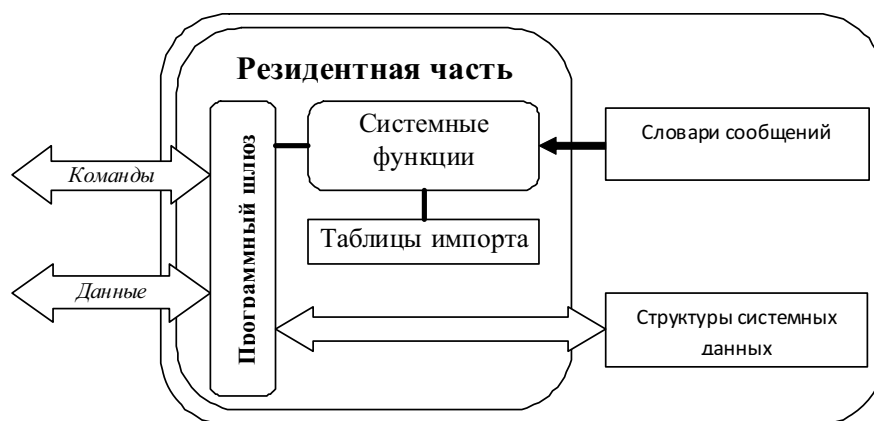


Рис. 8.2. Структурная схема подключаемого модуля

Основной управляющий элемент плагина – программный шлюз, обеспечивающий корректные вызовы зарегистрированных системных функций, динамическое выделение требуемого объема оперативной памяти для хранения системных структур данных, передачу результатов выполнения команд и возврат управления ядру системы после завершения выполнения функции, соответствующей данному плагину.

При помощи таблиц импорта шлюз сообщает ядру во время подключения модуля о том, какие именно системные функции модуль может выполнять. По умолчанию при подключении модуля место в оперативной памяти для данных не резервируется. Загружается лишь резидентная часть, а программный шлюз, входящий в нее, создает соответствующие структуры данных в памяти по мере необходимости.

Еще одна особенность структуры плагина – это наличие специальных словарей сообщений. Отметим, что при статичной привязке сообщений к программному модулю требуется полностью изменять программный код модуля или хранить несколько конечных вариантов программы, использующих различные языки в своих сообщениях. В ИИС ИМСЛОГ все манипуляции сообщениями производятся через систему селекторов, которым сопоставляются сообщения из словаря, соответствующего выбранному языку общения с пользователем. Выбор языка и загрузка соответствующего словаря производится ядром системы во время начального запуска, что дает возможность не занимать оперативную память неиспользуемыми в данном сеансе работы компонентами плагина. Добавление новых словарей производится специальной программой перевода модулей и не затрагивает их программный код.

Множество команд ядра является базисом, необходимым для интеллектуального анализа данных и знаний на основе вышеприведенных методов тестового распознавания образов и выполнения ряда вспомогательных функций. Этот базис может быть дополнен с целью расширения возможностей ИИС ИМСЛОГ. При этом обеспечивается совместимость «сверху вниз». Это означает, что при обновлении ядра не требуется повторная инсталляция, и прикладные системы, построенные на основе более поздней версии ядра, могут использовать ранее созданные плагины без каких-либо ограничений.

Программная реализация. В процессе программной реализации ИИС ИМСЛОГ применялись идеи и принципы структурного, нелинейного и объектно-ориентированного программирования.

Конечный программный продукт (библиотека C++ классов, модуль-ядро, динамически подключаемые модули-плагины, базовый модуль интеллектуального интерфейса) получен с использованием системы программирования Borland C++ Builder, а также API (Application Programming Interface) и GUI (Graphics User Interface) операционной среды Windows. Для создания когнитивных

средств использовалась библиотека OpenGL, предоставляющая богатые возможности по рендерингу поверхностей и объектов в трехмерном пространстве. Взаимодействие с пользователем организовано в виде интеллектуального многооконного интерфейса и реализовано в MDI (Multi Document Interface) форме. Для обеспечения переносимости программного кода расширения языка C++ не использовались.

Библиотека C++ классов спроектирована с учетом особенностей разработанных методов тестового распознавания образов и содержит определения базовых объектов, которые могут динамически создаваться, обрабатываться и уничтожаться по мере необходимости. К базовым относятся математические объекты, используемые в алгоритмах (например, векторы и матрицы различных типов), структурные элементы базы знаний, а также компоненты создаваемой интеллектуальной системы. В частности, интеллектуальная система в целом тоже рассматривается как базовый объект.

Создание прикладных интеллектуальных систем. Процесс создания прикладной интеллектуальной системы анализа информации и принятия решений с использованием ИИС ИМСЛОГ включает четыре этапа [24].

На первом этапе осуществляется систематизация и структуризация информации, необходимой для решения задач в данной проблемной области, а также определяются функциональный состав системы и ее архитектура на основе выбора подходов и методов распознавания образов, наиболее подходящих для решения именно этих задач.

Второй этап состоит в компоновке требуемой конфигурации прикладной интеллектуальной системы путем подключения к ядру соответствующих программных модулей. При необходимости, на этом же этапе часть модулей может быть модифицирована и/или могут быть созданы новые модули с учетом специфики данной проблемной области.

На третьем этапе создается и анализируется база данных и знаний с целью построения решающих правил, используемых при распознавании предъявляемых для анализа объектов.

Четвертый этап – это окончательная настройка системы для передачи заказчику.

В зависимости от решаемых задач конструируются прикладные интеллектуальные распознающие системы различной конфигурации. Построенные системы являются открытыми в плане оперативного наращивания их функциональных возможностей по интеллектуальному анализу информации и поддержке принятия решений с применением новых подходов, алгоритмов и когнитивных средств.

Методологическая основа ИИС ИМСЛОГ. Используются методы тестового распознавания образов, ориентированные на матричную модель представления данных и знаний (гл. 3), алгоритмы выявления закономерностей, включая

построение логических тестов. Принятие решений основано на сочетании логико-комбинаторного и логико-вероятностного подходов к тестовому распознаванию образов с привлечением графических и когнитивных средств визуализации информационных структур.

8.2.1. Программная реализация методов определения весовых коэффициентов в ИИС ИМСЛОГ

Рассмотренные в главе 4 методы и подход к определению весовых коэффициентов признаков были реализованы в виде трех функциональных модулей (плагинов) инструментального средства ИМСЛОГ в среде разработки Borland C++ Builder.

Каждый плагин представляет собой библиотеку динамической компоновки (Dynamic Link Library – DLL).

В основе построения каждого плагина лежит базовый класс `gakPlugIn`, реализующий набор виртуальных функций программного интерфейса плагина с ядром. От базового класса порождается класс-потомок, наследующий интерфейсные методы и имеющий собственные методы, реализующие конкретные алгоритмы. Методы класса-потомка регистрируются при помощи вызовов функции регистрации и после этого становится возможен их вызов через ядро.

Каждый плагин содержит функцию регистрации `RegistryProc`, предназначенную для начальной инициализации плагина, по средствам функций:

`RegisterPublishedMethod` – устанавливает краткое описание плагина и метод, который будет вызван при запуске этого модуля на исполнение;

`RegisterInputEntry` – устанавливает тип и описание входного значения. Количество вызовов этой функции внутри `RegistryProc` соответствует числу входов;

`RegisterOutputEntry` – устанавливает тип и описание выходного значения. Количество вызовов этой функции внутри `RegistryProc` соответствует числу выходов;

`RegisterSignature` – подпись плагина, используется во избежание конфликтных ситуаций между модулями различных версий.

Созданные модули имеют три входных значения (матрицы Q , R , T) и два выходных: вектор существенных признаков (целочисленный) и вектор весовых коэффициентов (вещественный).

В связи с тем, что подход, основанный на мультимножествах, содержит три метода, которые могут быть использованы как по отдельности, так и в совокупности, для данного подхода было добавлено диалоговое окно, дающее пользователю возможность выбора.

Для того чтобы созданный функциональный модуль подключить к ИС ИМСЛОГ необходимо добавить его в файл конфигурации системы с помощью специальной программы-конфигуратора (рис 8.3).

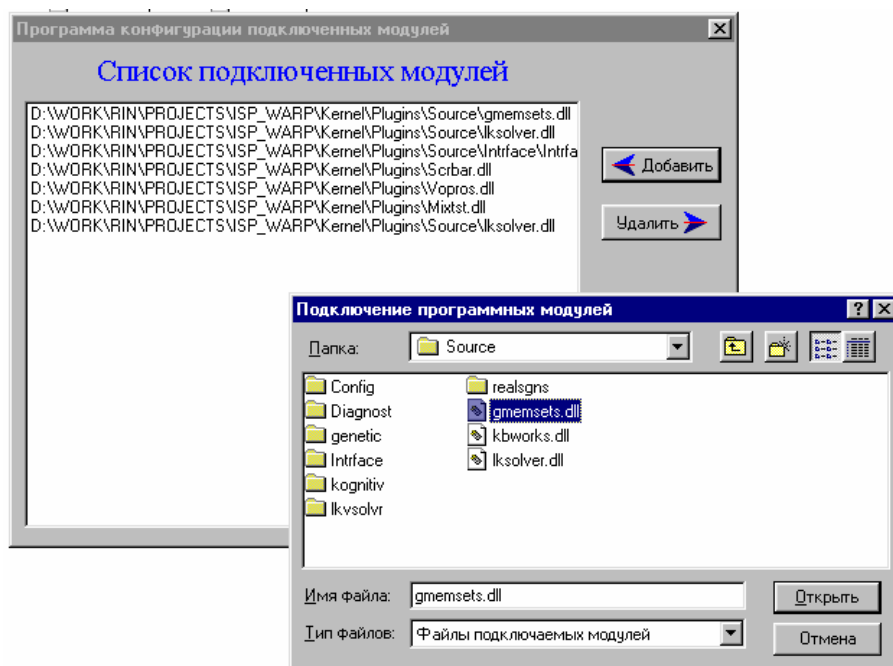


Рис. 8.3. Подключение программного модуля

При запуске инструментального средства ИМСЛОГ ядро через конфигурационный файл получает информацию о том, какие модули должны быть подключены к системе. Затем у каждого из этих модулей запрашивается функция RegistryProc. За графическое представление модулей отвечает модуль интерфейса пользователя.

8.3. Методология распознавания состояний динамических объектов в интеллектуальной системе IReDSS

Анализ ПО показал, что в перечисленных системах, во-первых, нет ориентации на реальное время (отклик в технических системах должен быть соизмерим с временным интервалом получения информации о состоянии ДО); во-вторых, в рамках ВР рассматривается только либо разложение его на составляющие (в т.ч., фильтрация), либо задача идентификации (восстановления зависимости), и не решается поставленная в этой работе задача распознавания состояний ДО. Программный комплекс IReDDS³² (рис. 8.4, 8.5) не только позволя-

³² Программная разработка А.Н. Мертвцова в рамках проекта РФФИ № 09-01-99014 р-офи. Сайт программного комплекса доступен по адресу: <http://redss.tom.ru>

ет без вмешательства в исходный код расширять базу используемых алгоритмов, но и решает задачи предоставления пользователю инструментария для обработки реального или модельного временного ряда; выбора модели из существующих прецедентов; использования АлР, адаптированных к предметной области, возможности их комбинирования; применения алгоритмов анализа ВР на основе методов частично наблюдаемых случайных процессов; обработки, анализа и представления результатов.

В отличие от традиционных ИИС, подразделяющихся на системы, решающие задачи анализа, и на системы, решающие задачи синтеза, система IReDSS объединяет в себе и способность к анализу (преобразование и интерпретация данных, диагностика, поддержка принятия решения) и к синтезу (управление), и может быть использована для обучения, мониторинга, прогнозирования, принимая исходные данные в виде временных рядов, или в виде матриц знаний (значения объект-признак).

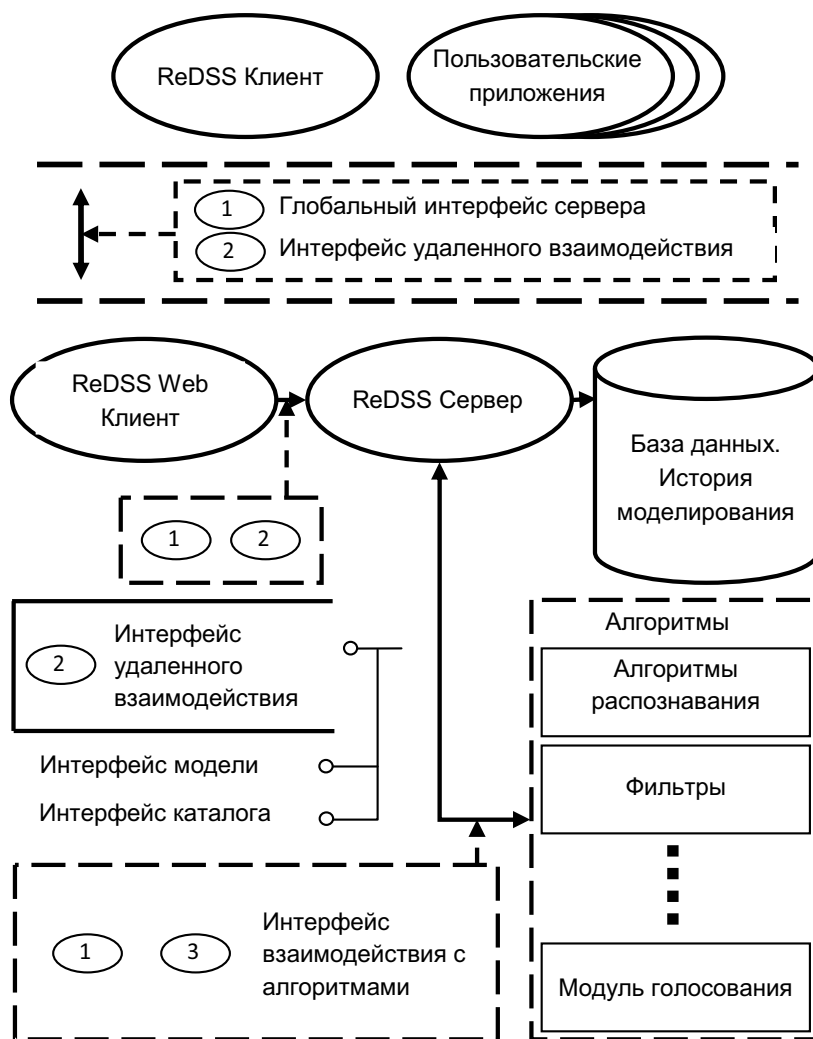


Рис. 8.4. Блок-схема программного комплекса

Программный комплекс IReDSS, реализующий методы и реализующие их алгоритмы распознавания состояний ДО [83, 146], имеет распределенную архитектуру. На стороне сервера, развертывается база данных (в которой хранится история моделирования), множество АлР, выполненных в виде подключаемых модулей, в состав комплекса входит WEB-клиент. К серверу можно подключатся из сторонних приложений, для этого доступны сборки с интерфейсами 1 и 2, с помощью кото-

рых можно добавить свой модуль (отдельную библиотеку) в каталог алгоритмов.

Модуль имеет параметры (типы входных и выходных данных). Тип модуля может быть произвольным и задается его разработчиком (распознаватель, фильтр, генератор ряда, модуль голосования и т.д.). Данные о конфигурации модулей, параметры, передаются серверу в виде модели (рис. 8.5). Модель представляет собой коллекцию ячеек, каждая из которых содержит информацию о блоке модели и его связи. Каждому подключенному к серверу клиенту соответствует свой экземпляр модели. Любые изменения в структуре модели, выполненные пользователем в клиентском приложении, отражаются на экземпляре, созданном на сервере. Установив необходимые значения параметров каждому модулю, обучив модули, требующие обучения, пользователь может начать процесс моделирования.

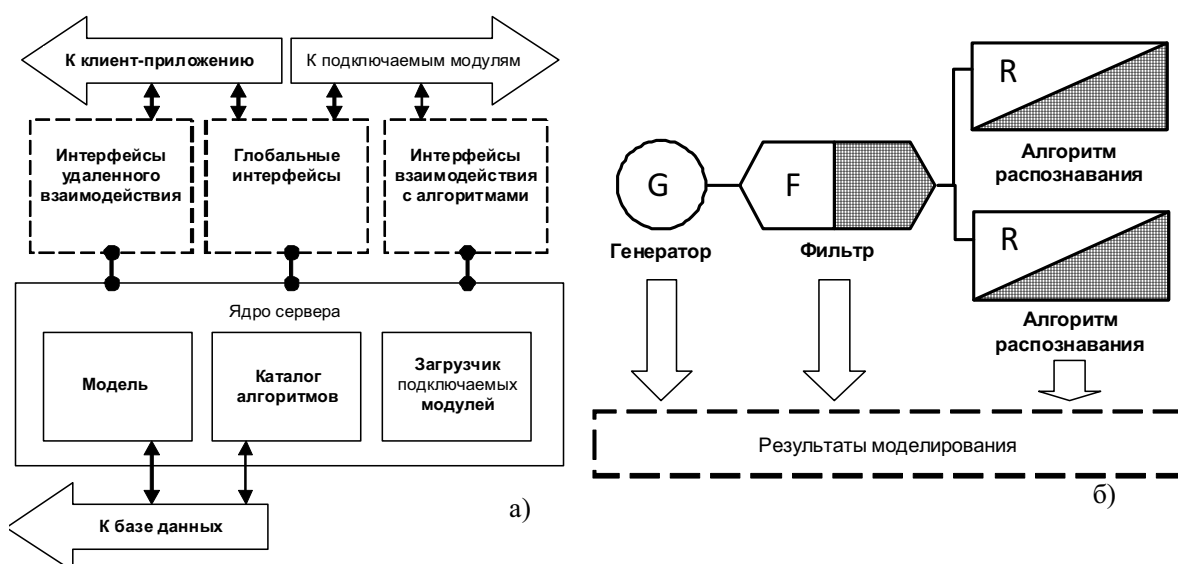


Рис. 8.5. Блок-схема сервера а) модель с описанием объектов и связей между ними, каталог алгоритмов с информацией о подключенных модулях; б) пример модели: результаты работы модулей сохраняется в базе данных

Программная реализация комплекса. Для реализации комплекса была выбрана платформа Microsoft.NET, имеющая объектно-ориентированную архитектуру, для поддержки распределенных приложений - Microsoft.NET Remoting³³, серверным доменом приложения выбрана служба Windows.

Методологическая основа ИИС IReDSS. Используются методы теории информации, теории вероятностей и математической статистики, теории распознавания образов (в том числе, тестового), методы нелинейной динамики, новые

³³ Маклин С., Нафтел Дж., Уильяме К. M15 Microsoft .NET Remoting / Пер. с англ. М.: Издательско-торговый дом «Русская Редакция», 2003. – 384 с.

алгоритмы выявления закономерностей на основе обобщенных эталонов, моделей авторегрессии и разностных схем, моделей прямой экстраполяции. Принятие решений основано на коллективном использовании логико-комбинаторных правил и информационно-вероятностных.

Выводы. Представлены две информационные и программные модели на основе математического аппарата гл.2-7 с целью решения слабоформализованных задач. Наряду с известной по ряду успешно решенных задач ИИС ИМСЛОГ представлена новая распределенная информационная система IReDDS для автоматизированной обработки ВР как модели сложного динамического объекта, функционирующего в условиях неопределенностей (параметрической, динамической), математическая модель которого является неточной (или неизвестной).

На основе IReDDS продемонстрирована работоспособность отдельных моделей распознавания состояний динамического объекта за реальное время с большой размерностью исходного признакового пространства. На конкретных прикладных задачах на ее основе изучено поведение и свойства семейств алгоритмов распознавания при различных вариациях интенсивностей шумов и величин параметров оптимизации, а именно: исследовано время и эффективность распознавания при изменении интервалов дискретизации, в условиях разных методов выбора системы информативных признаков, видов распределения и дисперсии шумов.

Результаты, представленные в 8-й главе, опубликованы в работах [20, 24, 145, 146, 174, 175].

«Вместо устойчивости и гармонии мы видим повсюду, куда ни обращаем свой взор, эволюционные процессы, приводящие ко все большему разнообразию и все возрастающей сложности» /И.Р. Пригожин/

Глава 9. РЕШЕНИЕ ПРИКЛАДНЫХ ЗАДАЧ

В главе рассматриваются четыре задачи, в которых использование разработанных выше подходов и реализующих их методов и алгоритмов распознавания состояний ДО (главы 2-5), реконструкции зашумленных динамических рядов является существенными в их решении: 1) задача построения модели наблюдателя при адаптивном управлении электромеханическим объектом; 2) задача экологического диагностирования и прогнозирования; 3) задача моделирования состояния здоровья пациента с бронхиальной астмой для оценки и прогноза его состояния; 4) задача предсказания неизвестных значений атрибутов в базах данных на примере данных АИС³⁴ «Торфяные ресурсы» и «Химия торфов» [62, 63].

9.1. Прикладная задача построения модели наблюдателя при адаптивном управлении электромеханическим объектом

В последнее время наметилось устойчивое расширение области применения частотно-регулируемых асинхронных приводов в различных промышленных механизмах [61, 69]. Это обусловлено многими факторами, в том числе снижением потребления энергии при внедрении таких электроприводов. В теории обобщенного электромеханического преобразователя получены уравнения динамики асинхронных двигателей (АД), однако их использование для построения динамических наблюдателей состояния и отработки новых алгоритмов управления (робастное осторожное управление [134], синергетическое управление [70] и др.) наталкивается на ряд трудностей, связанных с недостаточной разработанностью методов идентификации объектов на основе систем нелинейных дифференциальных уравнений.

Отличительной привлекательной чертой аналитического конструирования агрегированных регуляторов (АКАР) является не только совмещение решения задач управления и идентификации параметров в соответствии с процедурой указанного метода, но и полностью аналитически построенные процедуры синтеза наблюдателей координат состояния, параметров ДО и внешних возмущений. Решение задач синтеза высокого уровня сложности практически недоступно известным методам существующей теории управления [105, 106, 158].

³⁴ Задача решалась в рамках проекта РФФИ 97-05-64022-а «Исследование процессов трансформации органического вещества торфов» (Л.И. Инишева, О.Г. Савичева, С.И. Пяткова (Колесникова), Т.В. Дементьева)

Согласно АКАР, на первом этапе, считая, что все переменные состояния доступны измерению, предлагается процедура синтеза векторного регулятора, обеспечивающего выполнение некоторых требований в форме инвариантов. Так, для ЭМС АД имеются два канала управления, которые позволяют сформировать следующие инварианты: технологический – заданную частоту вращения вала, и электромагнитный – постоянство потокосцепления ротора. Другими словами, целью управления является поддержание заданной частоты вращения при оптимальном состоянии магнитной цепи АД.

Синтез регулятора выполняется с использованием расширенной модели АЭП, представленной в системе координат, вращающейся с угловой частотой, равной частоте вращения поля ротора и ориентированной по направлению вектора потокосцепления ротора.

Процедура синтеза динамического наблюдателя для АД разбивается на два этапа. На первом этапе предполагаем, что в объекте с течением времени могут изменяться такие параметры, как активные сопротивления статора и ротора, а также момент статической нагрузки на валу двигателя. Тогда, в соответствии с методом синтеза динамических регуляторов, вводится первая совокупность многообразий, на которой обеспечивается наблюдение заданных параметров АД. При этом предлагается два способа синтеза наблюдателей: при первом способе не требуется построения дополнительных наблюдателей производных, но в тоже время предполагает вычисление интегралов в аналитическом виде, что не всегда возможно; процедура построения наблюдателя вторым способом значительно проще, но в реальных системах могут сказываться ошибки вычисления используемых в ней производных измеряемых координат.

Приводится пример построения наблюдателя для синтеза адаптивного управления для нелинейной (в общем виде) стохастической системы на основе знания о скорости изменения измеряемых переменных. Для построения производных координат используются две модели (гл. 2, алгоритмы 2.7 и 2.8) с целью исследования на основе имитационного моделирования их предпочтительности.

9.1.1. Задача управления и синтеза наблюдателя состояния нелинейного электромеханического объекта на основе модели состояния

Математическая модель АД и постановка задачи. При определенных допущениях модель динамического движения АД может быть представлена следующей системой уравнений [61, 69]:

$$\begin{aligned}
\frac{di_{sx}(t)}{dt} &= \frac{1}{L_e} \left[U_{sx}(t) - R_e i_{sx}(t) + K_r A_r \psi_{rx}(t) + K_r z_p \omega_r(t) \psi_{ry}(t) \right] + \omega_k(t) i_{sy}(t), \\
\frac{di_{sy}(t)}{dt} &= \frac{1}{L_e} \left[U_{sy}(t) - R_e i_{sy}(t) + K_r A_r \psi_{ry}(t) - K_r z_p \omega_r(t) \psi_{rx}(t) \right] + \omega_k(t) i_{sx}(t), \\
\frac{d\psi_{rx}(t)}{dt} &= R_r K_r i_{sx}(t) - A_r \psi_{rx}(t) + \omega_k(t) \psi_{ry}(t) - z_p \omega_r \psi_{ry}(t), \\
\frac{d\psi_{ry}(t)}{dt} &= R_r K_r i_{sy}(t) - A_r \psi_{ry}(t) - \omega_k(t) \psi_{rx}(t) + z_p \omega_r \psi_{rx}(t), \\
\frac{d\omega_r(t)}{dt} &= \frac{1}{[J_r + J_c(t)]} [M_{\mathcal{E}}(t) - M_c(t)], \quad \frac{d\theta_r(t)}{dt} = \omega_r(t), \\
M_{\mathcal{E}}(t) &= \frac{3}{2} K_r z_p [\psi_{rx}(t) i_{sy}(t) - \psi_{ry}(t) i_{sx}(t)],
\end{aligned} \tag{9.1}$$

где $L_e = L_s - \frac{L_m^2}{L_r}$; $R_e = R_s + (R_r K_r^2)$; $K_r = \frac{L_m}{L_r}$; $A_r = \frac{R_r}{L_r}$;

В данных уравнениях приняты следующие обозначения: индексы « rx », « ry », « sx », « sy » использованы для обозначения координатных осей ротора и статора, соответственно; R_s , L_s' – сопротивление и индуктивность рассеяния статора; R_r , L_r' – приведенные значения сопротивления и индуктивности рассеяния ротора; L_m – взаимоиנדуктивность обмоток статора и ротора; L_s , L_r – полные индуктивности статора и приведенная индуктивность ротора, $L_s = L_s' + L_m$, $L_r = L_r' + L_m$; $U_{sx}(t)$, $i_{sx}(t)$, $U_{sy}(t)$, $i_{sy}(t)$ – напряжения и токи статора по координатным осям x и y ; $\psi_{rx}(t)$, $\psi_{ry}(t)$ – приведенные потокосцепления ротора по осям x и y ; $\omega_r(t)$ – механическая угловая частота ротора; $M_{\mathcal{E}}(t)$ – электромагнитный момент двигателя; z_p – число пар полюсов; J_r – момент инерции ротора двигателя; $J_c(t)$ – момент инерции механической нагрузки, приведенный к валу двигателя; $M_c(t)$ – механический момент сопротивления на валу двигателя; $\theta_r(t)$ – механическое угловое положение ротора.

Рассмотрим условие согласованности пространственного расположения координатных осей $\{x, y\}$ с ориентацией потокосцеплений $\psi_{rx}(t)$ и $\psi_{ry}(t)$, позволяющее сократить число переменных состояний в исходных уравнениях:

$$\psi_{ry}(t) = 0, \quad \frac{d\psi_{ry}(t)}{dt} = 0, \quad \psi_{rx}(t) = \psi_r(t), \tag{9.2}$$

Подставим условия (9.2) в уравнения (9.1), 4-е уравнение из (9.1) примет вид:

$$0 = R_r K_r i_{sy}(t) - \omega_k(t) \psi_r(t) + z_p \omega_r \psi_r(t), \tag{9.3}$$

Выразим из уравнения (9.3) переменную $\omega_k(t)$, подставим ее в уравнения (9.1), где эта переменная присутствует.

В дальнейшем, для организации системы управления АД текущее значение угла $\theta_k(t)$ поворота координатной системы $\{x, y\}$ вычисляется в соответствии с уравнением:

$$\frac{d\theta_k(t)}{dt} = \frac{K_r R_r i_{sy}(t)}{\psi_r(t)} + z_p \omega_r(t).$$

Принимая во внимание следующие обозначения: $x(t) = [x_1(t), x_2(t), x_3(t), x_4(t), x_5(t)]^T = [i_{sx}(t), i_{sy}(t), \psi_r(t), \omega_r(t), \theta_r(t)]^T$ – вектор состояния АД; $U(t) = [U_{sx}(t), U_{sy}(t)]^T$ – вектор управляющих воздействий, систему уравнений (9.1) можно записать в виде:

$$\begin{aligned} \dot{x}_1(t) &= -a_1 x_1(t) + a_2 \frac{1}{x_3(t)} x_2^2(t) + a_3 x_3(t) + a_4 x_2(t) x_4(t) + a_5 U_{sx}(t), \\ \dot{x}_2(t) &= -a_2 x_2(t) - \frac{1}{x_3(t)} x_1(t) - a_6 x_2(t) - a_7 x_4(t) x_3(t) - a_4 x_1(t) x_4(t) + a_5 U_{sy}(t), \\ \dot{x}_3(t) &= a_2 x_1(t) - a_8 x_3(t), \\ \dot{x}_4(t) &= a_9 a_{10}(t) x_3(t) x_2(t) - a_{10}(t) M_c(t), \end{aligned} \quad (9.4)$$

где $a_1 = \frac{R_e}{L_e}$; $a_2 = K_r R_r$; $a_3 = \frac{K_r A_r}{L_e}$; $a_4 = z_p$; $a_5 = \frac{1}{L_e}$; $a_6 = \frac{R_e}{L_e}$; $a_7 = \frac{K_r z_p}{L_e}$; $a_8 = A_r$;
 $a_9 = \frac{3}{2} z_p K_r$; $a_{10}(t) = \frac{1}{J_r + J_c(t)}$.

Ставится задача синтеза управления $u \in R^2$, как функции координат состояния, обеспечивающего асимптотическую устойчивость замкнутой системы относительно статического состояния, связанного условиями выполнения заданных режимов, а также построения наблюдателя для переменной $\omega_r(t)$.

9.1.2. Синтез непрерывного векторного закона управления асинхронным двигателем

Согласно этапам метода АКАР выберем систему инвариант:

–технологический инвариант – стабилизация частоты вращения вала двигателя:

$$x_4 - x_{40} = 0; \quad (9.5)$$

–энергетический инвариант – поддержание статического соотношения между магнитным потоком, частотой вращения ротора и моментом на валу двигателя, которое необходимо поддерживать в работе:

$$x_3^2 - x_{30}^2 = 0. \quad (9.6)$$

Следуя АКАР при векторном управлении сначала используется параллельная совокупность инвариантных многообразий ψ (равенство нулю которых $\psi=0$ обеспечивает достижение поставленных целей, или инвариантов), а затем осуществляется последовательный перевод изображающей точки от одного притягивающего многообразия к другому в соответствии с процедурой синтеза скалярного управления.

На первом этапе синтеза введем первую параллельную совокупность инвариантных многообразий (верхний индекс указывает номер этапа синтеза):

$$\begin{aligned} \psi_1^{(1)} &= x_1 - \varphi_1(x_3, x_4), \\ \psi_2^{(1)} &= x_2 - \varphi_2(x_3, x_4). \end{aligned} \quad (9.7)$$

Макропеременные $\psi_1^{(1)}, \psi_2^{(1)}$, должны удовлетворять основным функциональным уравнениям:

$$\begin{aligned} \omega_1^{(1)} \dot{\psi}_1^{(1)} + \psi_1^{(1)} &= 0, \\ \omega_2^{(1)} \dot{\psi}_2^{(1)} + \psi_2^{(1)} &= 0, \end{aligned} \quad (9.8)$$

которые при $\omega_1^{(1)} > 0$, $\omega_2^{(1)} > 0$ являются уравнениями Эйлера-Лагранжа относительно устойчивых экстремалей, доставляющих минимум на траекториях движения замкнутой системы сопровождающему оптимизирующему функционалу вида:

$$\Phi = \int_0^{\infty} \sum_{i=1}^2 \left((\omega_i^{(1)})^2 (\dot{\psi}_i^{(1)})^2 + (\psi_i^{(1)})^2 \right) dt, \quad i = 1, 2.$$

Весовые коэффициенты $\omega_1^{(1)}$, $\omega_2^{(1)}$ определяют задаваемое время движения изображающей точки системы до пересечения многообразий и являются параметрами настройки регулятора.

Подставим выражение (9.7) в уравнения (9.8) и с учетом (9.4), получаем:

$$\begin{aligned} U_{sx} &= \frac{1}{a_5} \left(a_1 x_1 - a_2 \frac{1}{x_3} x_2^2 - a_3 x_3 - a_4 x_2 x_4 + \frac{\partial \varphi_1}{\partial x_3} (a_2 x_1 - a_8 x_3) + \frac{\partial \varphi_1}{\partial x_4} (a_9 a_{10} x_3 x_2 - a_{10} M_c) - \frac{x_1 - \varphi_1}{T_1^{(1)}} \right), \\ U_{sy} &= \frac{1}{a_5} \left(a_2 x_2 \frac{1}{x_3} x_1 + a_6 x_2 + a_7 x_4 x_3 + a_4 x_1 x_4 + \frac{\partial \varphi_2}{\partial x_3} (a_2 x_1 - a_8 x_3) + \frac{\partial \varphi_2}{\partial x_4} (a_9 a_{10} x_3 x_2 - a_{10} M_c) - \frac{x_2 - \varphi_2}{T_2^{(1)}} \right). \end{aligned} \quad (9.9)$$

В результате попадания изображающей точки на параллельную совокупность многообразий $\psi_i^{(1)} = 0, i = 1, 2$ происходит динамическая декомпозиция замкнутой системы «АД – регулятор» (9.4), содержательно интерпретируемая как класс решений, которые удовлетворяют заданным целевым многообразиям (9.7). Согласно формулам (9.4), (9.8), (9.9), динамика системы (9.4) на пересечении многообразий $\psi_i^{(1)} = 0$ описывается декомпозированной моделью:

$$\begin{aligned}\dot{x}_3 &= a_2\varphi_1 - a_8x_3, \\ \dot{x}_4 &= a_9a_{10}x_3\varphi_2 - a_{10}M_c.\end{aligned}\tag{9.10}$$

Управления (9.9) в АКАР интерпретируются как внешние управляющие воздействия. Величины φ_1, φ_2 , подлежащие определению, интерпретируются как внутренние управляющие воздействия для декомпозированной системы уравнений (9.10).

На втором этапе синтеза вводится совокупность инвариантных многообразий для внутреннего управления, требования к которым сформулированы инвариантами (9.5, 9.6):

$$\begin{aligned}\psi_1^{(2)} &= x_3^2 - x_{30}^2, \\ \psi_2^{(2)} &= x_4 - x_{40}\end{aligned}\tag{9.11}$$

Макропеременные $\psi_1^{(2)}, \psi_2^{(2)}$ должны удовлетворять основным функциональным уравнениям:

$$\begin{aligned}\omega_1^{(2)}\dot{\psi}_1^{(2)} + \psi_1^{(2)} &= 0, \\ \omega_2^{(2)}\dot{\psi}_2^{(2)} + \psi_2^{(2)} &= 0,\end{aligned}\tag{9.12}$$

которые при $\omega_1^{(2)} > 0, \omega_2^{(2)} > 0$, являются уравнениями Эйлера-Лагранжа относительно устойчивых экстремалей, доставляющих минимум на траекториях движения замкнутой системы (9.10) сопровождающему оптимизирующему функционалу вида:

$$\Phi = \int_0^{\infty} \sum_{i=1}^2 \left((\omega_i^{(2)})^2 (\dot{\psi}_i^{(2)})^2 + (\psi_i^{(2)})^2 \right) dt.$$

Весовые коэффициенты $T_1^{(2)}, T_2^{(2)}$ определяют задаваемое время движения ИТ системы (9.10) до пересечения многообразий и являются параметрами настройки регулятора.

Подставим выражение (9.11) в уравнения (9.12) и с учетом (9.10), получаем:

$$\varphi_1 = \frac{(2\omega_1^{(2)}a_8 - 1)x_3^2 + x_{30}^2}{2x_3\omega_1^{(2)}a_2}, \quad \varphi_2 = \frac{a_{10}M_c\omega_2^{(2)} - (x_4 - x_{40})}{\omega_2^{(2)}a_9a_{10}x_3}.\tag{9.13}$$

Найдем частные производные от φ_1 и φ_2 :

$$\begin{aligned}\frac{\partial \varphi_1}{\partial x_3} &= \frac{a_8}{a_2} - \frac{1}{2\omega_1^{(2)}a_2} - \frac{x_{30}^2}{2\omega_1^{(2)}a_2x_3^2}, \quad \frac{\partial \varphi_1}{\partial x_4} = 0, \\ \frac{\partial \varphi_2}{\partial x_3} &= -\frac{a_{10}M_c\omega_2^{(2)} - (x_4 - x_{40})}{\omega_2^{(2)}a_9a_{10}x_3^2}, \quad \frac{\partial \varphi_2}{\partial x_4} = -\frac{1}{\omega_2^{(2)}a_9a_{10}x_3}.\end{aligned}\tag{9.14}$$

Подставив выражения (9.13), (9.14) в выражения для управлений (9.9), получим искомые управляющие воздействия, которые вместе с уравнениями (9.4) образуют регулятор для системы (9.1):

$$U_{sx} = \frac{1}{a_5} \left(a_1 x_1 - a_2 \frac{1}{x_3} x_2^2 - a_3 x_3 - a_4 x_2 x_4 + \left(\frac{a_8}{a_2} - \frac{1}{2\omega_1^{(2)} a_2} - \frac{x_{30}^2}{2\omega_1^{(2)} a_2 x_3^2} \right) (a_2 x_1 - a_8 x_3) \right) - \frac{1}{a_5} \left(\frac{x_1}{\omega_1^{(1)}} - \frac{(2\omega_1^{(2)} a_8 - 1) x_3^2 + x_{30}^2}{2a_2 x_3 \omega_1^{(1)} \omega_1^{(2)}} \right), \quad (9.15)$$

$$U_{sy} = \frac{1}{a_5} \left(a_2 x_2 \frac{1}{x_3} x_1 + a_6 x_2 + a_7 x_4 x_3 + a_4 x_1 x_4 - (a_2 x_1 - a_8 x_3) \frac{a_{10} M_c \omega_2^{(2)} - (x_4 - x_{40})}{\omega_2^{(2)} a_9 a_{10} x_3^2} \right) - \frac{1}{a_5} \left(\frac{a_9 a_{10} x_3 x_2 - a_{10} M_c}{\omega_2^2 a_9 a_{10} x_3} + \frac{x_2}{\omega_2^{(1)}} - \frac{a_{10} M_c \omega_2^{(2)} + (x_4 - x_{40})}{a_9 a_{10} \omega_2^{(1)} \omega_2^{(2)} x_3} \right).$$

9.1.3. Конструирование наблюдателя частоты вращения ротора асинхронным двигателем

Рассмотрим задачу синтеза наблюдателя для оценки фазовой переменной $x_4(t)$ (частота $\omega_r(t)$ вращения ротора) в условиях стационарности параметров $J_c(t) = J_c$ и $M_c(t) = M_c$ механической нагрузки АД.

Используя метод АКАР [70], по числу оцениваемых переменных состояния $x_4(t)$ введем макропеременную $\psi(t)$

$$\psi(t) = \phi(x_1, x_2, x_3, x_4) - \hat{\phi}(t), \quad (9.16)$$

удовлетворяющую решению однородного дифференциального уравнения:

$$L^{-1} \dot{\psi}(t) - \psi(t) = 0, \quad (9.17)$$

где L – параметр оптимизации АКАР.

Одним из вариантов выбора промежуточной функции $\phi(x_1, x_2, x_3, x_4)$ является суперпозиция измеряемых $x_1(t)$, $x_2(t)$, $x_3(t)$ и неизмеряемой, $x_4(t)$ фазовых переменных системы:

$$\phi(x_1, x_2, x_3, x_4) = x_1(t) + x_2(t) + x_3(t) + x_4(t). \quad (9.18)$$

При этом предполагается, что функция $\hat{\phi}(t)$ является асимптотической оценкой функции $\phi(x_1, x_2, x_3, x_4)$. Подставляя в уравнение (9.17) выражение (9.6) с учетом выражения (9.18) получим дифференциальное уравнение:

$$\frac{\partial \phi}{\partial x_1} \dot{x}_1(t) + \frac{\partial \phi}{\partial x_2} \dot{x}_2(t) + \frac{\partial \phi}{\partial x_3} \dot{x}_3(t) + \frac{\partial \phi}{\partial x_4} \dot{x}_4(t) - \frac{d\hat{\phi}(t)}{dt} - L\phi(t) + L\hat{\phi}(t) = 0 \quad (9.19)$$

С целью исключения неизмеряемой фазовой переменной из уравнения (9.19), дополнительно введем функции $\gamma(t)$, $\Gamma_i(t)$, $i = 1, 2, 3$, обеспечивающие выполнение следующего равенства:

$$\begin{aligned} & \frac{\partial \phi}{\partial x_1} \dot{x}_1(t) + \frac{\partial \phi}{\partial x_2} \dot{x}_2(t) + \frac{\partial \phi}{\partial x_3} \dot{x}_3(t) + \frac{\partial \phi}{\partial x_4} \dot{x}_4(t) - L\phi(t) - \\ & - \Gamma_1(t) \dot{x}_1(t) - \Gamma_2(t) \dot{x}_2(t) - \Gamma_3(t) \dot{x}_3(t) - \gamma(t) = 0 \end{aligned} \quad (9.20)$$

Используя равенства (9.16)– (9.19) и выражение (9.20) уравнение (9.39) запишем в виде:

$$\begin{aligned} & \dot{x}_1(t) + \dot{x}_2(t) + \dot{x}_3(t) + \dot{x}_4(t) - L(x_1(t) + x_2(t) + x_3(t) + x_4(t)) - \\ & - \Gamma_1(t) \dot{x}_1(t) - \Gamma_2(t) \dot{x}_2(t) - \Gamma_3(t) \dot{x}_3(t) - \gamma(t) = 0, \text{ или} \\ & \dot{x}_1(1 - \Gamma_1) + \dot{x}_2(1 - \Gamma_2) + \dot{x}_3(1 - \Gamma_3) + \dot{x}_4 - \\ & - L(x_1 + x_2 + x_3 + x_4) - \gamma = 0 \end{aligned} \quad (9.21)$$

После подстановки уравнений (9.4), являющихся моделью ЭМО, в равенство (9.41) запишем:

$$\begin{aligned} & (-a_1 x_1 + a_2 x_3^{-1} x_2^2 + a_3 x_3 + a_4 x_2 x_4 + a_5 U_{sx})(1 - \Gamma_1) + \\ & + (-a_2 x_3^{-1} x_2 x_1 - a_6 x_2 - a_7 x_4 x_3 - a_4 x_1 x_4 + a_5 U_{sy})(1 - \Gamma_2) + (1 - \Gamma_3)(a_2 x_1 - a_8 x_3) + \\ & + a_9 a_{10} x_3 x_2 - a_{10} M_C - L(x_1 + x_2 + x_3 + x_4) - \gamma = 0. \end{aligned} \quad (9.22)$$

Функции $\gamma(t)$, $\Gamma_i(t)$ выбираются таким образом, чтобы исключить из уравнений (9.22) неизмеряемую фазовую переменную $x_4(t)$. Одним из вариантов выбора данных функций является вариант исключения из системы уравнений (9.22) всех слагаемых, содержащих управляющие воздействия за счет функции $\gamma(t)$ вида:

$$\gamma(t) = a_5 [U_{sx}(t)(1 - \Gamma_1(t)) + U_{sy}(t)(1 - \Gamma_2(t))], \quad (9.23)$$

После подстановки выражения (9.23) для функции $\gamma(t)$ в уравнение (9.22) запишем:

$$\begin{aligned} & (-a_1 x_1 + a_2 x_3^{-1} x_2^2 + a_3 x_3 + a_4 x_2 x_4)(1 - \Gamma_1) + \\ & + (-a_2 x_3^{-1} x_2 x_1 - a_6 x_2 - a_7 x_4 x_3 - a_4 x_1 x_4)(1 - \Gamma_2) + \\ & + (1 - \Gamma_3)(a_2 x_1 - a_8 x_3) + a_9 a_{10} x_3 x_2 - a_{10} M_C - L(x_1 + x_2 + x_3 + x_4) = 0. \end{aligned} \quad (9.24)$$

Для определения функций $\Gamma_i(t)$, $i = 1, 2, 3$ в уравнении (9.24) приравняем к нулю выражения, являющиеся множителями переменных x_4^0 , x_4 .

$$\begin{aligned}
x_4 : a_4 x_2 (1 - \Gamma_1) - (a_7 x_3 + a_4 x_1) (1 - \Gamma_2) - L &= 0, \\
x_4^0 : (-a_1 x_1 + a_2 x_3^{-1} x_2^2 + a_3 x_3) (1 - \Gamma_1) - (a_2 x_3^{-1} x_2 x_1 + a_6 x_2) (1 - \Gamma_2) + & (9.25) \\
+ (a_2 x_1 - a_8 x_3) \tilde{\Gamma}_3 + a_9 a_{10} x_3 x_2 - a_{10} M_C - L(x_1 + x_2 + x_3) &= 0.
\end{aligned}$$

Из системы (9.25) находим функции Γ_1 и Γ_3 :

$$\Gamma_1 = 1 - \frac{L}{a_4 x_2}, \quad \Gamma_2 = 0,$$

$$\Gamma_3 = 1 + \frac{-a_1 x_1 + a_2 x_3^{-1} x_2^2 + a_3 x_3}{a_2 x_1 - a_8 x_3} \frac{L}{a_4 x_2} + \frac{a_9 a_{10} x_3 x_2 - a_{10} M_C - L(x_1 + x_2 + x_3)}{a_2 x_1 - a_8 x_3}.$$

Тогда уравнение (9.20) с учетом найденных коэффициентов γ , $\Gamma_1, \Gamma_2, \Gamma_3$, может быть переписано в виде:

$$\begin{aligned}
\Gamma_1 \dot{x}_1(t) + \Gamma_2 \dot{x}_2(t) + \Gamma_3 \dot{x}_3(t) - \dot{\hat{\phi}}(t) + \gamma + L \hat{\phi}(t) &= 0, \text{ или} \\
\dot{\hat{\phi}}(t) = \Gamma_1 \dot{x}_1(t) + \Gamma_2 \dot{x}_2(t) + \Gamma_3 \dot{x}_3(t) + \gamma + L \hat{\phi}(t) &= 0. \quad (9.26)
\end{aligned}$$

Решение дифференциального уравнения (9.26) $\hat{\phi}(t)$ будет представлять собой асимптотическую оценку переменной $\phi(t)$, выбранной в виде $\phi = x_1 + x_2 + x_3 + x_4$. Тогда искомая оценка для переменной x_4 будет иметь вид:

$$x_4 = \hat{\phi}(t) - x_1 - x_2 - x_3.$$

Представленный вариант наблюдателя для переменной x_4 связан с решением задачи построения дополнительных наблюдателей производных измеряемых переменных $x_1(t), x_2(t), x_3(t)$ и в разрешении уравнения (9.26) относительно $\hat{\phi}(t)$. Ниже представлены иллюстрации (рис. 9.1-9.3) управления по АКАР и наблюдателей, полученных согласно структурным схемам процедур оценивания производной по алгоритмам 2.7, 2.8.

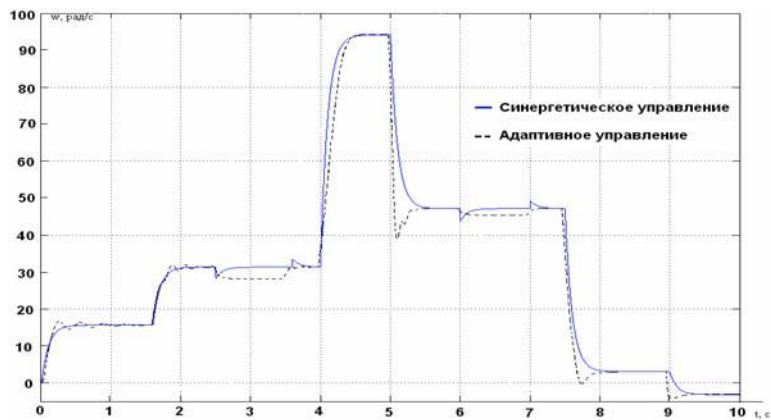


Рис. 9.1. Управление по модификации метода АКАР и адаптивному по эталонной модели

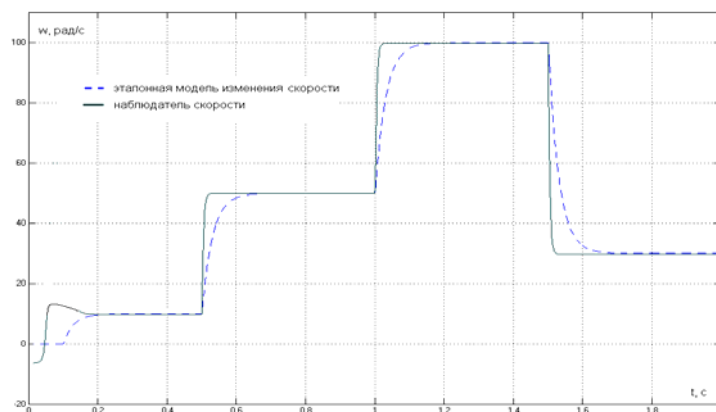


Рис. 9.2. Наблюдатель скорости вращения ротора с производной, определяемой по алгоритму 2.7 (на основе нелинейной регрессии)

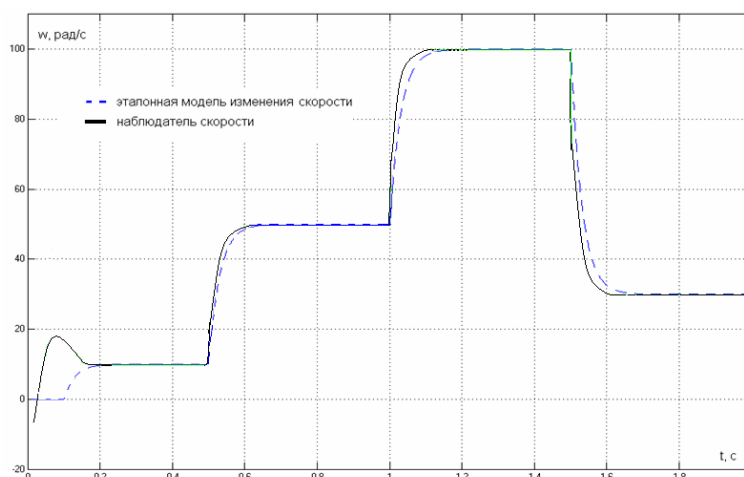


Рис. 9.3. Наблюдатель скорости вращения ротора с производной, определяемой по алгоритму 2.8 (на основе ARADS)

9.2. Прикладная задача экологического диагностирования и прогнозирования на основе временных рядов и обобщенных эталонов

В работе [122] приведены результаты разработки теоретических основ ведения экологических измерений, классификации природных объектов и оценки скорости развития контаминации (в частности, событий, приводящих к необратимым последствиям) в пространстве.

Целью данного раздела является, во-первых, формализация задачи оценивания состояний экологического объекта как задачи распознавания состояний динамического объекта; во-вторых, создание подхода и реализующих его простых и эффективных алгоритмов для автоматизированного выявления закономерностей в признаковом экологическом пространстве, для принятия итогового

решения о принадлежности состояния экологического объекта одному из выделенных (классов, образов) и прогнозирования состояния объекта во времени.

Экологический объект обладает всеми атрибутами сложного ДО (далее ЭДО), поскольку характеризуется свойствами: а) объект не имеет полного аналитического описания и полностью описывающей его модели (слабо формализован); б) образует во времени, изменяя характеристики (показатели, признаки), многомерный нестационарный случайный процесс (при дискретизации – слабо-предсказуемый временной ряд); в) временной ряд (в общем) имеет нелинейную модель тренда.

Алгоритмизация распознавания состояний ЭДО, модель характеристик которого задана в виде многомерного временного ряда, позволит проводить автоматизированное выделение природных объектов по их описанию, осуществлять решение задачи экологического районирования, оценивать степень экологической напряженности и «глубину» экологического прогнозирования, а также проводить оценивание скорости изменения качества природных ресурсов и объектов.

Основные понятия, определения, обозначения. Далее используется понятие «экологическое пространство» $E=R \times F \times T$, где R - реальное подпространство с координатами (x, y, z) и $F=(f_1, f_2, \dots, f_M)$ - факторное (параметрическое) подпространство, f_k – косвенный признак, f_{jk} - значение k -го признака для j -го экологического объекта a_j ($j = \overline{1, N}$, $k = \overline{1, M}$, $t = \overline{0, T}$, N – общее число объектов, M -число косвенных признаков).

Декартово произведение $E=A \times F \times T$ множества реальных объектов $A=\{a_j\}$, множества показателей (факторов) F и множества моментов времени T называют экологическим полигоном $G \in E=R \times F \times t$, $G=\{x_{jkt}\}$ ($j = \overline{1, N}$, N – общее число объектов, $k = \overline{1, M}$, M -число косвенных признаков). Элементом x_{jkt} этого произведения является значение k -го показателя на j -ом объекте в момент времени t . Ставится задача о выделении природных объектов и экологических аномалий (почвы, воды, воздуха) [90, 122].

Обозначим векторный временной ряд (ВР) через $x_j=\{x_{jkt}\}$ ($k = \overline{1, M}$, $t = \overline{0, T}$) (динамика развития k -го показателя (параметра) j -го объекта во времени).

Поставим каждому ВР x_j в соответствие ряды d_j , размеченные номерами градаций показателя, или возможных его состояний (например, не\удовлетворительное, хорошее, отличное, закодированное целочисленным значением) [60, 90]:

$$d_{jkt} = \begin{cases} l, & x_{jkt} \in \Omega_l^k, l \in \{1, \dots, I_k\}, \\ 0, & x_{jkt} \notin \Omega_l^k, \forall l = \overline{1, I_k}. \end{cases}$$

Замечание. Для кодирования показателя, принимающего действительные значения в некотором диапазоне, интервал допустимых значений показателя f_k

разбивают на участки с требуемой точностью и каждый участок кодируют целым числом. Отметим, что разметку ВР интервалами роста и падения значений показателя на данном объекте [60] можно интерпретировать как сегментацию ВР в смысле 2-х состояний.

Определим функцию сходства временных рядов x_i, x_j (как моделей экологических объектов a_i, a_j , соответственно) по k -му показателю $\lambda_k(x_i, x_j)$ в виде:

$$\lambda_k(x_i, x_j) = \sum_{t=1}^T \chi_{jkt}, \quad \chi_{jkt} = \begin{cases} 1, & d_{jkt} = d_{ikt}, d_{ikt} \neq 0, d_{jkt} \neq 0, \\ 0, & \text{иначе.} \end{cases}$$

В частном случае, когда разметка ВР затруднена, в качестве функции сходства временных рядов x_i, x_j можно взять значение $\lambda_k(x_i, x_j) = \omega_k \sqrt{\sum_{t=1}^T (x_{ikt} - x_{jkt})^2}$,

где ω_k - весовой коэффициент k -го показателя.

Определение 9.1. Временной ряд является размеченным, если каждому его элементу поставлена в соответствие метка одного из состояний $B = \{1, \dots, l\}$; в общем случае, знак из алфавита B , интерпретируемый как некоторое качество показателя (увеличение, уменьшение и пр.).

Определение 9.2. Два показателя (признака, меры сходства и т.д) C_1 и C_2 независимы по предпочтению от других показателей C_3, \dots, C_n , если предпочтения между альтернативами, различающимися лишь оценками по C_1 и C_2 , не зависят от фиксированных значений по другим показателям.

В предположении независимости показателей по предпочтению функцию сходства временных рядов x_i, x_j по всему множеству показателей (параметров) определим как аддитивную весовую свертку значений независимых показателей:

$$\lambda_{ij} = \lambda(x_i, x_j) = \sum_{k=1}^M \alpha_k \lambda_k(x_i, x_j), \quad \sum_{k=1}^M \alpha_k = 1 .$$

Таким образом, исходная информация о полигоне как совокупности значений «объект-показатель-время» представлена многомерными временными рядами, и далее становится возможным проведение классификации ЭДО или выделение групп похожих объектов относительно заданного показателя, либо относительно заданного набора показателей.

Понятие обучающей выборки введено в п.1.2, закономерность будем понимать в смысле определения 1.16.

Состояние ЭДО – участок (фрагмент) многомерного ВР, удовлетворяющего определенным ограничениям. Наблюдаемый объект – это участок (фрагмент) временного ряда или весь временной ряд. Применительно к распознаванию состояний ЭДО постановка задачи будет следующая.

Постановка задачи. Пусть Ω_0 - множество состояний реального ЭДО, Ω_Y - множество реализаций непрерывного случайного процесса изменения состояний ЭДО, $\Omega_Y \subseteq \Omega_0$, Ω_{TS} - подмножество дискретных наблюдений реализаций состояний ЭДО (обучающих прецедентов, доступных измерению временных рядов), $\Omega_{TS} \subseteq \Omega_Y$, $\Omega = \{\Omega_1, \Omega_2, \dots, \Omega_l\}$ - множество образов (классов состояний ЭДО, выделенных в соответствии с целью исследования), $\Omega \subseteq \Omega_Y$. Будем предполагать: 1) существует отображение $f: \Omega_0 \rightarrow \Omega_Y$ - математическая модель состояний реального ДО; 2) существует (неизвестное, однозначное) отображение $g: \Omega_0 \rightarrow \Omega_{TS}$ - дискретная модель состояний ДО (реальные значения состояний ДО отображаются в дискретные подмножества $\Omega_i, i = \overline{1, l}$).

По заданной обучающей выборке - множеству значений

$\Theta_Y = \{y_j^l, \rho_j^l\}_{j=1}^N \subset \Omega_{TS}$, где y_j^l - j -й объект (фрагмент ВР), ρ_j^l - метка номера со-

стояния ЭДО, $l = \overline{1, k}$ - число серий наблюдений, - требуется: 1) построить алгоритм распознавания $A(\Theta_Y): \Omega_{TS} \rightarrow \Omega$ (возможно, множество алгоритмов), осуще-

ствляющему оценку метки состояния ЭДО $d \in I^\Omega = \{1, \dots, l\}$ по совокупности но-

вых наблюдений (k -окну) $Y_j^k = (y_{j1}, \dots, y_{jk}), Y_j^k \in \Omega_{TS}, k \geq 1, j \in T^L = \{1, \dots, T-k\}$; 2) синтезировать итоговое решающее правило, согласующее частные решения в случае нескольких решающих правил; 3) оценить прогностические возможности моделирования по обобщенным эталонам состояний ЭДО в смысле минимума выбранного критерия качества (скользящий контроль).

Подход к решению задачи. Кратко отметим общую направленность исследований в свете вышеизложенной информационной модели исследования сложных ДО.

1. *Формирование эталонов* (образов) экологических объектов по обучающей выборке и оценки их типичности в виде весовых коэффициентов g -эталонов на основе алгоритма 2.1.

2. *Коррекция информационных эталонов* при пополнении обучающей выборки выполняется следующим образом. С учетом веса эталонов осуществляется проверка на близость к эталонам разных классов. Если добавляемый объект находится ближе к эталону своего класса, то выполняется корректировка значений признаков этого эталона, вес его также увеличивается. Если добавляемый объект находится ближе к эталону не своего класса, то он объявляется новым

эталоном с единичным весом (рис. 2.4, с)) (алгоритм 2.1). Эксперименты показывают хорошую устойчивость к шумам и дополнительной информации, которая в случае типичности добавляемых объектов только увеличивает вес уже существующих g -эталонов.

3. *Принятие решения* об отнесении нового экологического объекта к одному из назначенных классов осуществляется по одному из метрических алгоритмов классификации (в том числе, предложенных выше - алгоритмы 2.2, 3.1): Для повышения надежности решения целесообразно применение FRiS-функции (как относительной меры сходства [16]), корректно работающей и в случае разной структуры классов (разного среднего внутриклассового расстояния) и коллективных правил на основе ММАИ (алгоритм 2.3, подробно гл. 3).

4. *Прогнозирование предельных состояний ЭДО*. Обозначим совокупность наблюдаемых временных рядов, соотнесенных с динамическими стохастическими экологическими объектами через (x_1, \dots, x_N) , $x_j = (x_{jk1}, \dots, x_{jkT})$, $j = \overline{1, N}$, $k = \overline{1, M}$. Сформируем ВР, состоящие из разностей соседних значений исходных ВР [5, 121, 166]:

$$z_j = (z_{j1}, \dots, z_{j, T-1}), z_{jt} = \sqrt{\sum_{k=1}^M (x_{jkt} - x_{jk, t+1})^2}, j = \overline{1, N}, t = \overline{t_0, T-1} \quad (9.27)$$

Способ формирования одномерного временного ряда, составленного из разностей соседних значений, близок к функции межточечных расстояний [167], удобной для исследований непуассоновских точечных потоков. Подобные временные ряды применялись в работе [121] для исследования свойств фазовых траекторий динамических систем. Формируя авторегрессионную модель по значениям ряда (9.27) (AR-модель) и рассматривая характеристики, в частности, AR-коэффициенты $a(t)$ и корни характеристического полинома $q(t)$ можно показать следующее [121]:

1) функции $a(t)$ и $q(t)$ являются убывающими, начиная с некоторого момента времени t для временных рядов (или экологических объектов) x_{jkt} , стремящихся к состоянию равновесия;

2) функция $a(t)$ стремится к некоторому постоянному значению, максимальный по модулю корень характеристического полинома $q(t)$ стремится к единице (начиная с некоторого момента времени t) для временных рядов x_{jkt} , отвечающих периодическому изменению свойств; дальнейшее исследование может быть связано с уточнением величины периода (например, на основе нижеизложенного алгоритма реконструкции фрагмента временного ряда);

3) функция $a(t)$ хаотично изменяется во времени, а максимальный по модулю корень характеристического полинома $q(t)$ может иметь разный характер: больше единицы для x_{jkt} (не отвечающим состояниям равновесия), меньше 1, равна 1; в этом случае однозначный прогноз невозможен.

Перечисленные прогностические признаки могут оказаться полезны при исследовании экологического объекта в динамике развития. Исследование на основе AR-модели (9.27) выявляет устойчивые равновесия, периодические движения и аттрактор, отвечающий «отрицательному образу» ЭДО, а также их области притяжения. Эти данные позволяют выявить основные типы эволюции ЭДО (восстановление удовлетворительного состояния, хроническое неудовлетворительное состояние, рецидивирующее плохое состояние и необратимо плохое состояние) и влиять на варианты их осуществления.

5. Реконструкция функциональной зависимости изменения показателей (признаков) ЭДО во времени на основе алгоритма 2.5. Знание динамической модели каждого состояния ЭДО позволит осуществлять краткосрочные прогнозы.

6. Иллюстративный пример моделирования на ПЭВМ. Численное моделирование проведено на реальных примерах временных рядов, содержащих результаты измерения некоторых характеристик торфов [62, 63] во времени.

Пример 9.1. В таблице 9.1 приведены результаты моделирования на основе алгоритма 2.5. по данным [62, 63]. (κ_1 , κ_2) - минимальное и максимальное содержание; Y_{av} / Y^* - среднее и прогнозное содержание со средним среднеквадратическим отклонением, равным 0.08.

Таблица 9.1. Результаты оценивания и прогнозирования значений некоторых показателей состава торфов низинного типа

Группа, вид торфа	Зольность, %		Сумма гуминовых кислот	
	(κ_1 , κ_2)	Y_{av} / Y^*	(κ_1 , κ_2)	Y_{av} / Y^*
Древесная группа	6,7–16,6	11,4 / 10,8	17,8–44,2	26,6/ 24,9
Древесно-травяной	8,8–15,5	11,6/ 11,1	16,9–37,8	24,3/ 21,7
Древесно-осоковый	10,5–15,5	12,8/ 12,1	20,8–37,8	26,2/ 24,3
Травяная группа	8,8–13,8	10,4/ 9,9	16,9–31,8	22,3/ 21,8
Травяной	3,4–16,5	8,2/ 9,8	16,5–36,3	25,5/ 25,8
Вахтовый	4,2–13,3	8,0/ 10,1	28,2–36,3	32,3/ 31,1
Осоковый	3,4–12,7	9,1/ 9,7	19,2–28,5	23,3/ 20,8
Травяно-моховая группа	4,2–16,5	8,0/ 9,6	16,5–32,4	23,9/ 22,6
(осоково-гипновый)	7,3–17,9	10,4/ 8,9	18,3–27,9	21,5/ 20,4
Моховая группа (гипновый)	4,7–14,5	8,1/ 9,1	17,9–27,3	22,2/ 21,5

Выводы. В разделе рассмотрен подход к оцениванию состояний экологического объекта как сложного динамического объекта, информация о котором доступна в виде нестационарного временного ряда. Исследование состояний объекта основано на построении обобщенных эталонов состояний объекта, од-

новременном применении методов прямой экстраполяции и распознавания образов для реконструкции функциональной зависимости изменения показателей (признаков) ЭДО во времени, анализе поведения коэффициентов и корней характеристического полинома соответствующей авторегрессионной зависимости.

9.3. Прикладная задача моделирования состояния здоровья пациента с бронхиальной астмой

В настоящее время осваивается универсальный показатель состояния основных функций человека: физической, психологической, социальной и духовной. Он получил наименование качество жизни (КЖ)³⁵. КЖ, связанное со здоровьем, позволяет изучить влияние заболевания и лечения на показатели КЖ больного человека, оценивая все составляющие здоровья. Понятие КЖ положено в основу понимания болезни и определения эффективных методов ее лечения. Слежение КЖ является общепринятым в международной практике высокоинформативным, чувствительным и экономичным методом оценки состояния здоровья как населения в целом, так и отдельных социальных групп.

Одним из средств оценки КЖ является мониторинг состояния здоровья, как населения в целом, так и отдельных его групп. Исследование КЖ актуально, в частности, в пульмонологии, пациентов с бронхиальной астмой (БА). При этом мониторинг состояния здоровья таких пациентов необходим для достижения и поддержания клинического контроля над болезнью. Инструментами обеспечения контроля над БА является: выявление пациентов с неконтролируемой БА; принятие своевременного решения о необходимости назначения или изменения терапии.

Задача о технологической поддержке мониторинга состояния здоровья пациента с БА средствами Web-технологии требует создания математической модели объекта мониторинга и выбора базирующегося на ней алгоритма динамической оценки состояния здоровья пациента с выдачей рекомендаций.

Постановка задачи. Представим население Ω кортежем подмножеств: $\Omega = \{\Omega_i\}$, где Ω_i – группы населения с разными состояниями здоровья, $i=0,1,\dots,l$. Например Ω_0 – трудоспособное население, Ω_1 – пенсионеры, Ω_2 – дети и т.д.; Ω_l – группа практически здоровых людей. Для упрощения задачи на этом этапе примем, что $\Omega_i \cap \Omega_j = \emptyset$. Далее введем факторное пространство F с координатами S_1, S_2, \dots, S_m (показателями состояния здоровья и симптомами), точки которого отражают состояние здоровья человека $y_j, j = 1, 2, \dots, n$. Разобьем интервалы возможных значений координат этого пространства на C_l смежных интервалов, соответствующих международным среднепопуляционным показателям КЖ со-

³⁵ World Health Organization. Quality of life group. What is it Quality of life? Wld. Hth. Forum. - 1996. - V. 1. - P. 29

стояния здоровья населения. Все пространство состояний здоровья будет иметь $\prod_{l=1}^m C_l$ ячеек, каждая из которых соответствует тому или иному среднепопуляционному показателю состояния здоровья человека, сопоставленного с группами $\Omega_i, i=0,1,\dots,I$.

Поставим задачу построения механизма распознавания состояния здоровья пациента по конкретным данным о симптомах и значениях измеряемых показателей физического состояния пациента, с учетом их возможной взаимной зависимости и погрешности измерений.

Моделирование процессов болезни и диагностирования проведем на примере проблемной области пульмонологии, а именно на примере мониторинга состояния здоровья пациентов с БА³⁶.

Подходы к решению задачи. Кратко остановимся на сути четырех подходов с разными моделями представления исходных данных и знаний.

Модель³⁷ 1. Для оценки состояния здоровья конкретного пациента с БА будем использовать следующие медицинские симптомы и показатели:

Таблица 9.2. Медицинские симптомы и показатели

1. Симптомы	2. Показатели
а) затрудненное дыхание; б) свистящие хрипы; в) кашель; г) чувство заложенности в груди; д) одышка.	а) ОФВ ₁ – объем форсированного выдоха за первую секунду; б) ФЖЕЛ – форсированная жизненная емкость легких; в) ПСВ – пиковая скорость выдоха; г) суточная вариабельность ПСВ; д) МОС – мгновенная объемная скорость при выдохе от ФЖЕЛ в %; е) ПОС – пиковая объемная скорость вдоха; ж) СОС – средняя объемная скорость форсированного выдоха.

Симптомы $C_l \in C = \{C_1, C_2, \dots, C_n\}$ могут принимать значения $[0; 1]$. Если симптом присутствует, то $C_l=1$, в противном случае $C_l=0$. Показатели $K_k \in K = \{K_1, K_2, \dots, K_m\}$ могут принимать численные значения от 0 до 1000. Необ-

³⁶ Качество жизни у больных бронхиальной астмой и хроническим обструктивным заболеванием легких / Под ред. А.Г. Чучалина. - Москва: Атмосфера, 2004. – 256 с.

³⁷ С.И. Колесникова, В.А. Шапцев, М.В. Якунина. Модель оценки состояния здоровья пациента с бронхиальной астмой. - Тюменский государственный университет.

ходимо учитывать погрешность при измерении показателей: $\delta = \sqrt{(m-1)^{-1} \sum_{k=1}^m (K_m - K_k)^2}$,

где $K_{эт}$ – эталонная величина показателя.

Изменение симптомов и показателей происходит во времени.

Для определения принадлежности пациентов к определенной группе используем методы кластеризации. Под кластеризацией понимается процесс разбиения множества состояний пациентов на классы, при котором элементы, объединенные в один класс, имеют большее сходство, чем элементы, принадлежащие разным классам (сравнение состояний при формировании кластеров должно вестись по атрибутам – симптомам и показателям (табл. 9.2, рис. 9.5)).

Пример³⁸ 9.2. На рис. 9.4 графически интерпретированы градации двух состояний пациентов: удовлетворительного (Normal) и неудовлетворительного (Pathological Process).

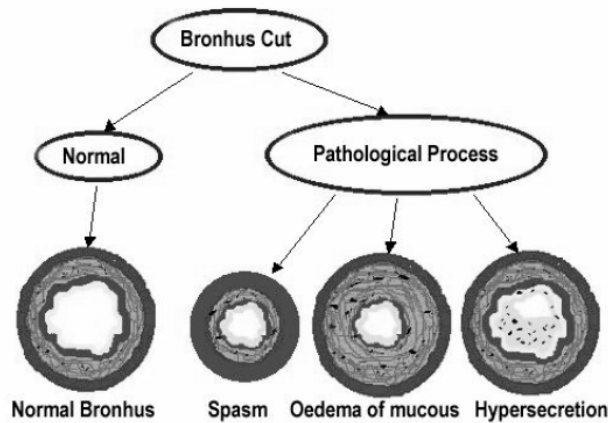


Рис.9.4. Три структуры патологических процессов БА

Разобьем пространство F_1 состояний здоровья пациентов с БА на три кластера (рис. 9.5): k_1 - неконтролируемая БА; k_2 - частично контролируемая БА; k_3 - контролируемая БА.

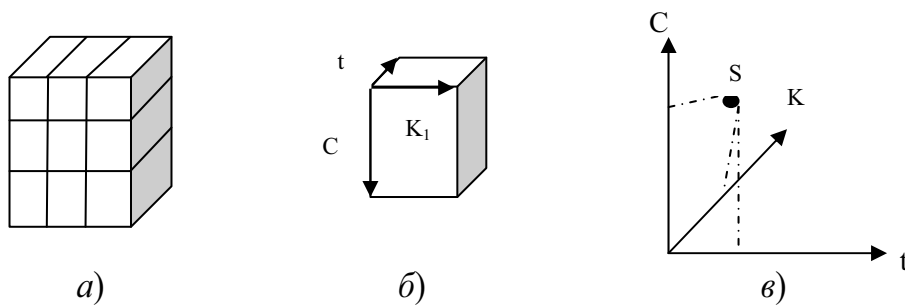


Рис. 9.5. а) - пространство состояний F_1 , б) – ячейка, в) - координата S в пространстве F_1

³⁸ Рисунок из работы: Янковская А.Е., Тетенев Ф.Ф., Черногорюк Г.Э. Отражение образного мышления специалиста в интеллектуальной распознающей системе патогенеза заболеваний (на примере бронхиальной астмы) // Новости искусственного интеллекта. – 1999. – № 1. – С. 129–140.

Пространство F_1 состояний здоровья пациентов с БА, можно геометрически интерпретировать в виде куба (рис. 9.5 а)), состоящего из ячеек (рис. 9.5 б)). Грани ячеек – это симптомы C_i , показатели K_k и моменты времени t . Координата $S=(C, K, t)$ пространства F_1 – это точка отображающая состояние здоровья конкретного пациента с БА y_j (рис. 9.5 с)).

Такому представлению могут соответствовать трехмерные матрицы данных и знаний типов: «симптом-показатель-время», «объект-показатель-время», «объект-симптом-время», и задача классификации заключается в нахождении близких (в определенном смысле) предварительно нормированных и центрированных временных рядов или фазовых траекторий разных динамических объектов (симптомов, пациентов). Выделение симптомов как динамических объектов связано, во-первых, с более медленной их динамикой по сравнению с показателями (признаками), во-вторых, с возможностью обнаружения новых закономерностей между симптомами и показателями.

Сопоставим модели «объект-показатель-время» временной ряд y_{ikt} , $k=1, m$ - значение k -го показателя на i -ом объекте в момент времени t . Сформируем одномерный ВР из многомерного, компоненты которого состоят из разностей соседних значений исходных предварительно нормированных и центрированных временных рядов [121, 167]:

$$z_j = (z_{j1}, \dots, z_{j, T-1}), z_{jt} = \sqrt{\sum_{k=1}^m \omega_{kt} (y_{jkt} - y_{jk, t+1})^2}, j = \overline{1, n}, t = \overline{t_0, T-1}. \quad (9.28)$$

Представление (9.28) обладает сглаживающим эффектом и позволяет сравнивать объекты (симптомы, пациенты) по изменению динамики всей совокупности показателей; ω_{kt} - весовые коэффициенты k -го показателя в момент времени t , определять которые можно по методу ММАИ (гл. 3), гарантирующему корректное взвешивание взаимозависимых альтернатив по многим критериям (или оценкам экспертов) в динамически изменяющихся наборах альтернатив.

Определим функцию сходства временных рядов y_i, y_j (как моделей классифицируемых объектов) по совокупности параметров $\lambda(x_i, x_j)$ в виде:

$$\lambda(y_i, y_j) = \sum_{t=1}^{T-1} \chi_{jlt}, \chi_{jlt} = \begin{cases} 1, & |z_{jt} - z_{it}| \leq \delta_t, \delta_t \geq 0, \\ 0, & \text{иначе.} \end{cases} \quad (9.29)$$

Здесь δ_t – параметр, задающий пороговое значение допуска на сходство объектов, зависящее от момента времени в силу нестационарности исходных временных рядов. Далее по матрице сходства, составленной на основании значений функции сходства (9.29) однозначно определяются группы объектов со сходным поведением [151].

Для идентификации связи между двумя показателями X, Y на основе ARADS и обучающей выборки с известными сочетаниями показаний наблю-

даемой Y и ненаблюдаемой X переменных (x_j, y_j) , $j = \overline{1, n}$, $x_j = X(j\Delta)$, $y_j = Y(j\Delta)$, $\Delta > 0$, получают динамические модели по алгоритму 2.5 $x = g(t)$, $y = f(t)$. В результате находят функциональную зависимость между наблюдаемой Y и ненаблюдаемой X переменными: $x = g(f^{-1}(t))$, где $f^{-1}(\cdot)$ - функция, обратная к $f(\cdot)$, значимость которой можно осуществить на дополнительной выборке.

Модель 2. Алгоритм для выявления закономерностей во временных рядах опирается на алгебраический подход [60, 139], развитый в гл. 2 для разметки стохастических ВР. Временной ряд является размеченным (определение 9.1), если каждому его элементу поставлен в соответствие знак из заданного алфавита A , например, $A = \{U, D, N\}$, где символы U и D интерпретируются как увеличение и уменьшение значения показателя данного объекта в данный момент времени, символ N соответствует отказу от классификации. В новом алгоритме 2.7. взят алфавит $A = \{v, v = 1, 2, \dots, n_f; N\}$ меток (номеров) функциональных зависимостей – претендентов из заданного класса f_v , $v = 1, n_f$.

Наряду с алгоритмом 2.7. приведем пример [60] алгоритма разметки для слабозашумленных ВР.

Пример 9.3. Каждому временному ряду $\{y_{ikt}\}$ ставятся в соответствие ряды

$\{u_{ikt}\}$ и $\{d_{ikt}\}$, где $u_{ikt} \in \{U, N\}$, $d_{ikt} \in \{D, N\}$, интерпретируемые как ряды, разме-

ченные интервалами роста и падения значений показателя (признака) на данном объекте. Здесь $\{y_{ikt}\}$ - декартово произведение множества объектов, множества признаков и множества моментов времени, элемент y_{ikt} - значение k -го показателя на i -ом объекте в момент времени t .

Разметка ВР $\{y_{ikt}\}$ осуществляется в соответствии с условием (аксиомой): элементу из последовательности $(y_{ikt_1}, \dots, y_{ikt_N})$ ставится в соответствие знак U , если выполнены соотношения:

$$\begin{aligned}
 & y_{ikt_1} < y_{ikt_2}, y_{ikt_{N-1}} < y_{ikt_N}, \\
 & d < \left((y_{ikt_N} - y_{ikt_1}) - \sum_{s \in T_D} (y_{ikt_s} - y_{ikt_{s+1}}) \right) (y_{ikt_N} - y_{ikt_1})^{-1}, \\
 & y_{ikt_1} < y_{ikt_s}, s = \overline{1, N}, \\
 & T_D = \{s : y_{ikt_{s+1}} < y_{ikt_s}\}, r \leq y_{ikt_N} - y_{ikt_1},
 \end{aligned} \tag{9.30}$$

где d, r – параметры процедуры разметки. Знак D ставится в соответствие элементу из последовательности $(y_{ikt_1}, \dots, y_{ikt_N})$, если выполнены вышеприведенные условия с учетом замены y_{ikt} на $(-y_{ikt} - \max_t y_{ikt})$. В противном случае ставится знак N .

Алгоритм для выявления закономерностей во временных рядах на основе его разметки содержит решение следующих подзадач (п. 2.4, алгоритм 2.7).

1. Разметка временного ряда по аксиомам для определения меры сходства ВР (расстояния между ВР) [44].

2. Построение матриц сходства между всеми парами временных рядов. Сходство (или расстояние) определяется числом размеченных интервалов совместного роста или спада значений признаков (показателей).

3. Определение групп объектов со сходным поведением для одного фиксированного показателя по матрице сходства.

4. Нахождение групп «похожих» объектов для произвольного набора показателей на основе задания матрицы смежности групп объектов и представления в виде множества графов [151].

Модель 3. Понятие n - m -полюсника (прямоугольник с n входами и m выходами) используется в таких проблемных областях, как радиотехника, проектирование логических устройств, компьютерные технологии и др. Предлагается³⁹ основанная на формализме n - m -полюсников методология для обоснования принятия диагностических решений в области медицины в прикладных интеллектуальных системах (рис. 9.6). База знаний для диагностики заболеваний лёгких дополняется описаниями синдромов⁴⁰ и симптомов.

На основе принимаемого (прикладной ИС) решения формируется структура из n - m -полюсников, предоставляющая полномасштабную системную картину диагностического процесса.

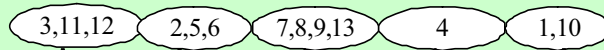
Начиная от первичных симптомов и их клинического обоснования, возможно продвижение к следующему уровню - синдромам, формирование основного синдрома и последующего заключения о возможных заболеваниях лёгких. Далее на основе симптомов дополнительного обследования принимается заключительный диагноз.

³⁹ Yankovskaya A.E, Galkin D.V., Chernogoryuk G.E. Computer visualization and cognitive graphics tools for applied intelligent systems // Proceedings of the IASTED International Conference on Automation, Control, and Information Technology (ACIT 2010), Juny 15-18, 2010 Novosibirsk, Russia, ACTA Press Anaheim/Calgary/Zurich, Vol. 1. – 2010 – pp. 249–253.

⁴⁰ Синдром – образ на основе базы знаний и экспертной оценки.

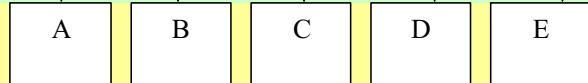
Первичные симптомы:

- 1) Кашель с мокротой
- 2) Общая слабость
- 3) Периодически повышение температуры до 38,0°
- 4) Одышка при умеренной физической нагрузке
- 5) Головная боль
- 6) Снижение аппетита
- 7) Укорочение перкуторного звука в подмышечной области справа
- 8) Бронхиальное дыхание под лопаткой справа
- 9) Усиление голосового дрожания над правым легким
- 10) Сухие низкие хрипы над правым легким
- 11) Ускорение СОЭ 42 мм/час
- 12) Количество лейкоцитов - $14 \times 10^9 / \text{л}$
- 13) При рентгенологическом обследовании - крупноочаговое затемнение в средней доле правого легкого

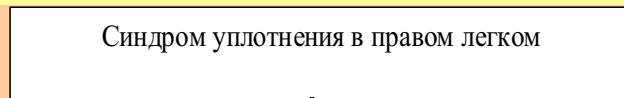


Синдромы:

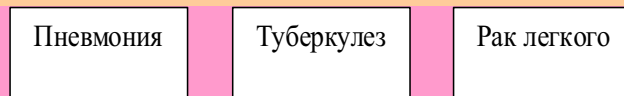
- A) Синдром системной воспалительной реакции
- B) Синдром интоксикации
- C) Синдром уплотнения в правом легком
- D) Синдром недостаточности дыхания
- E) Бронхитический синдром



Основной синдром:

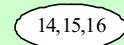


Возможные заболевания:



Симптомы дополнительного обследования:

- 14) Ателектаз средней доли правого легкого
- 15) Деформация среднедолевого бронха и его обструкция
- 16) Обнаружение раковых клеток в биоптате из правого среднедолевого бронха



Заключительный диагноз

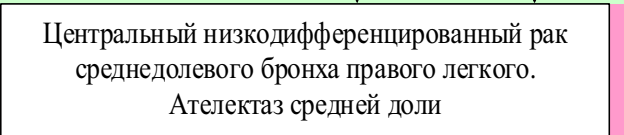


Рис.9.6. Представление знаний в медицинской диагностике на основе n-т-полюсников

На рис. 9.6 (в целях удобства восприятия) различные уровни принятия решения в диагностическом процессе окрашены в различный цвет, а входы каждого n - m -полюсника 1-го уровня представлены номерами симптомов, заключённых в овал, от которого исходит одна линия со стрелкой к соответствующему n - m -полюснику. Например, входами n - m -полюсника, отмеченного символом B , являются симптомы 2,5,6, расшифровка которых представлена слева на рис. 9.6: общая слабость, головная боль, снижение аппетита соответственно. Выходом n - m -полюсника, отмеченного символом B , является линия со стрелкой, сопоставленная синдрому интоксикации, представленного слева на рис. 9.6, и являющаяся 2-м входом n - m -полюсника «синдром уплотнения в правом лёгком (основной синдром)» и так далее до построения заключительного диагноза.

Сочетание натуралистического (имеющего отображение в обычной реальности) подхода и основанного на n - m -полюсниках принятия и обоснования принятия диагностических решений является основой создания эффективных когнитивных средств при разработке прикладных интеллектуальных систем.

Модель 4. Модель, используемая для представления многопризнаковых объектов, в том числе и с противоречивыми признаковыми описаниями. В работе [131] предложены методы упорядочения и классификации совокупности сложных объектов, которые базируются на теории метрических пространств мультимножеств (гл. 4, п. 4.2, пример приложение 4). Метод упорядочения объектов основан на оценке их близости по отношению к некоторому «идеальному» объекту (эталону) в многопризнаковом пространстве. Метод классификации объектов позволяет строить обобщенное решающее правило для их отбора, которое аппроксимирует различные, в том числе и противоречивые, правила экспертной сортировки объектов. Данный подход может быть применен с вынесением более высокого качества итоговых решений за счет изложенной в [46, 85] методики (гл. 2, п. 2.1) формирования обобщенных эталонов, адаптивной к обучающей выборке (а не назначаемой априорно) и за счет новых мер относительного сходства мультимножеств (4.11).

Выводы. Процессы, происходящие в организме больного, многомерны, подвержены взаимовлияниям, вступают во взаимодействия по типу разветвленной последовательной множественности. Представленные вычислительно реализуемые модели болезни и диагностирования, основанные на выявлении различного рода закономерностей во временных рядах, сопровождающих процесс болезни, сделают возможными: выявление новых связей между различными процессами с последующим выходом на патофизиологические эффекты; проведение экспериментов с вводом «возмущающих» элементов (лечебных воздействий); имитирование индивидуальной схемы развития заболевания у пациента с целью лечебной коррекции.

9.4. Прикладная задача предсказания неизвестных значений непрерывных атрибутов в базах данных

С точки зрения поиска закономерностей использование базы данных в качестве обучающего множества вызывает определенные трудности [10, 31], среди которых наиболее важными являются следующие: недостаточность доступной информации, необходимой для определения класса объекта; частичное отсутствие информации, ее «зашумленность» в некотором смысле и/или изменение информации во времени (например, стоимость акции, товара, значение химического (минералогического) состава исследуемого образца [62, 63] и т.д.); объем данных (примеров объектов) достаточно велик и растет, причем некоторые удачные в прошлом объекты теряют свое качество от времени из-за постоянного развития технологий; процесс оценки и ранжирования объектов занимает достаточно существенное время.

В работе [31] излагаются методы обнаружения знаний, использование которых частично позволяет решать задачу обобщения объектов в условиях определенных ограничений (например, независимости признаков, описывающих объект, или дискретности значений признаков).

Здесь рассматривается случай признаков с непрерывными значениями, и обсуждается вопрос об оценке пропущенных значений при обработке реальных массивов, связанных с химическим и минералогическим составом торфов [62, 63].

Постановка задачи. Пусть по базе данных как по обучающему множеству выделены классы. Требуется оценить отсутствующие значения признаков в выделенных K классах. При этом, не исключается возможность того, что размерность признакового пространства достаточно велика, а значения признаков могут зависеть от времени [10], т.е. характеристики динамического объекта (вида торфа) – упорядоченная совокупность значений атрибутов базы данных, т.к. в действительности признаки могут быть подвержены действию случайных возмущений и оставаться постоянными только в среднем.

Решение задачи. Для решения задачи в контексте рассматриваемых в книге подходов к распознаванию состояний ДО рассмотрим три варианта применения авторегрессионных моделей, используемых для предсказания будущих значений наблюдаемого ряда по настоящим и прошлым значениям: 1-й – на основе обобщенных эталонов и параметризации модели авторегрессии (п. 2.3), 2-й – на основе непараметрической нелинейной регрессионной модели (п. 2.5), 3-й излагается ниже.

Будем предполагать, что обрабатываемые атрибуты имеют непрерывные значения (случайные) из заданного интервала. Рассмотрим проблему использо-

вания объектов обучающей выборки для нахождения общих свойств выделенных I классов и использования этих свойств а) для предсказания неизвестных и/или потерянных значений; б) для распознавания объектов, не входящих в обучающую выборку.

Предположим, что в качестве математической модели значений атрибута x объектов одного класса Ω_i рассматривается линейная авторегрессия со случайными коэффициентами⁴¹, описываемая уравнениями:

$$x_k^i = a_1(k, i) \cdot x_{k-1}^i + \dots + a_p(k, i) \cdot x_{k-p}^i + \sigma_0(i) \cdot \xi_k^i, i = \overline{1, I}, \quad (9.31)$$

где $a_l(k, i) = a_l(i) + \sigma_l(i) \cdot \eta_l(k, i)$, $\{\eta_l(k, i)\}$, $\{\xi_k^i\}$, $l = \overline{1, p}$, $p = 1, 2, \dots$ – независимые последовательности независимых одинаково распределенных случайных величин с нулевым средним и единичной дисперсией, k – номер временного интервала (номер «замера» значения признака) ($k=1, 2, \dots$), $\sigma_l(i)$ – постоянные неотрицательные параметры. Отметим, что модель (9.31) допускает «шумы» двойкой природы без предположения их гауссовости.

Требуется оценить по наблюдениям признака $\{x_k^i\}$ вектор параметров $\theta(i) = (a_1(i), a_2(i), \dots, a_p(i))^T$, т.е. идентифицировать ДО с целью предсказания его поведения.

Вектор оцениваемых параметров $\theta(i) = (a_1(i), a_2(i), \dots, a_p(i))^T$ по сути отражает специфику класса Ω_i (п. 2.5). Представление наблюдений признака $\{x_k^i\}$ (9.31) равносильно следующему:

$$x_k^i = \theta^T \cdot X_{k-1}^i + \psi_k^i, \psi_k^i = \sum_{l=1}^p \eta_l(k, i) \cdot x_{k-l}^i + \xi_k^i, X_{k-1}^i = (x_{k-1}^i, \dots, x_{k-p}^i)^T.$$

Введем дополнительные обозначения:

$$\mathbf{A}_i = \begin{pmatrix} a_1(i) & \dots & a_p(i) \\ \mathbf{I}_{p-1} & \dots & 0 \end{pmatrix}, \mathbf{G}_k(i) = \begin{pmatrix} \eta_1(k, i) & \dots & \eta_p(k, i) \\ 0 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & 0 \end{pmatrix}, \mathbf{C}_i = \mathbf{E} \cdot \mathbf{G}_k(i) \otimes \mathbf{G}_k(i),$$

где \mathbf{I}_{p-1} – единичная матрица порядка $(p-1)$.

Теоретическим обоснованием ниже излагаемого алгоритма 9.1 является теорема⁴² 9.1.

⁴¹ Модель, развиваемая рядом научных школ, в частности, В.В. Коневым, профессором ТГУ, и его учениками.

⁴² Кашковский Д. В. Последовательная идентификация параметров авторегрессии со случайными коэффициентами // Вестник Томского гос. ун-та. – 2006. – № 293. – С. 105–109.

Теорема 9.1. Если уравнение (9.31) имеет стационарное решение, то последовательная оценка по методу наименьших квадратов $\theta^*(i) = (a_1^*(i), a_2^*(i), \dots, a_p^*(i))^T$ сходится в среднеквадратическом к истинному значению вектора параметров $\theta(i) = (a_1(i), a_2(i), \dots, a_p(i))^T$, а верхняя граница среднеквадратической точности вектора оценки $\theta^*(i)$ обратно пропорциональна длительности наблюдений.

Отметим, что при $\sigma_l(i) = 0, l = \overline{1, p}$ уравнение (9.31) имеет стационарное решение, если выполнены условия:

- 1) все корни характеристического уравнения $\lambda^p - a_1 \cdot \lambda^{p-1} - \dots - a_p = 0$ по модулю меньше единицы;
- 2) распределение ξ_k^i имеет положительную плотность в интервале $[-d, d]$ для некоторого $d > 0$;
- 3) абсолютные величины собственных значений матрицы $\mathbf{A} \otimes \mathbf{A} + \mathbf{C}$ меньше единицы ($\mathbf{A} \otimes \mathbf{A}$ – кронекерово произведение).

Алгоритм 9.1. Оценка неизвестного значения признака (атрибута)

Вход: наблюдения признака $\{x_k\}$.

Выход: оценка параметров $\theta^* = (a_1^*, a_2^*, \dots, a_p^*)^T$, оценка отсутствующего

значения признака x_k .

1. Задаем пороговое значение h .

2. Вычисляем момент остановки наблюдений признака x ,

$$\tau_h = \inf \left\{ n : \left\| M_n^{-2} \right\|^{1/2} \leq h^{-1} \right\}, \text{ где } M_n = \sum_{k=1}^n X_{k-1} X_{k-1}^T.$$

3. Определяем последовательную оценку $\theta^* = (a_1^*, a_2^*, \dots, a_p^*)^T$ вектора оцениваемых параметров модели $\theta = (a_1, a_2, \dots, a_p)^T$ по формуле:

$$\theta^* = \overline{M}_{\tau_h}^{-1} \cdot \sum_{k=1}^{\tau_h} b_k \cdot X_{k-1} x_k, \text{ где } \overline{M}_{\tau_h}^{-1} = \sum_{k=1}^{\tau_h} b_k \cdot X_{k-1} X_{k-1}^T, b_k = \begin{cases} 1, & k < \tau_h \\ \alpha_h, & k = \tau_h \end{cases}, \text{ величина}$$

на α_h удовлетворяет уравнению $h^{-1} = \left\| \left(\sum_{k=1}^{\tau_h-1} X_{k-1} X_{k-1}^T + \alpha_h X_{\tau_h-1} X_{\tau_h-1}^T \right)^{-2} \right\|$.

4. Вычисляем отсутствующее значение признака x по формуле:

$$x_k = a_1^*(k) \cdot x_{k-1} + \dots + a_p^*(k) \cdot x_{k-p}.$$

Вернуть: x_k .

Пример 9.4. Рассмотрим пример, где представлены численные значения оценок вектора параметров $\theta = (a_1, a_2)^T$, полученных в результате работы последовательной процедуры, основанной на методе наименьших квадратов, а также выборочные значения основных характеристик: h – количество наблюдений, τ_h – среднее $vS(\theta_1^*)$, $S(\theta_2^*)$ – (табл. 9.3) и табл. 9.3 значения свойств предложенной процедуры моделировался процесс (9.31) второго порядка:

$$x_k = a_1(k) \cdot x_{k-1} + a_2(k) \cdot x_{k-2} + \xi_k, \quad a_l(k) = a_l + \sigma_l(i) \cdot \eta_l(k), \quad l = 1, 2.$$

Таблица 9.3. Выборочные характеристики процедуры

a_1	a_2	h	τ_h	$S(\theta_{1st}^*)$	$S(\theta_1^*)$	$S(\theta_{2st}^*)$	$S(\theta_2^*)$
0.1	0.5	50	53.04	0.161	0.560	0.062	0.083
-0.45	0	200	271.9	0.063	0.426	0.026	0.096
0.3	0.3	200	273.2	0.091	0.342	0.002	0.081

Выводы. Для построения процедуры получения оценки неизвестного значения признака (атрибута) объектов, информация о которых хранится в базе данных, во-первых, изложен алгоритм 9.1, не требующий жестких ограничений на распределения шумов (такого как требование гауссовости); во-вторых, применен алгоритм 2.5 моделирования по временному ряду на основе обобщенных эталонов и параметризации модели авторегрессии. Моделирование показало незначительное различие в результатах (для стационарного процесса) с большим преимуществом в простоте и скорости реализации второго подхода, лишенного неопределенности в выборе порядка авторегрессии (при подгонке статистических данных, например, моделями ARCH(p) приходится обращаться к слишком большим значениям параметра p). При исследовании 3-х подходов к моделированию нестационарного процесса (по дисперсии шума) 1-й подход показал лучшие результаты на меньшем числе данных (размер окна наблюдения $k=20-25$).

Результаты, представленные в 9-й главе, кроме отмеченных по тексту, опубликованы в работах [24, 62, 63, 90, 91, 145, 146, 174].

ЗАКЛЮЧЕНИЕ

В книге представлен обзор методов и формализация задачи распознавания состояний сложного динамического объекта с целью применения в информационной и программной модели оценивания состояний сложного объекта на основе временных рядов.

Рассмотрены модели выявления закономерностей и соответствующие алгоритмы распознавания состояний сложного динамического объекта:

- модель на основе обобщенных эталонов состояний, дающая возможность решения трех задач: распознавания состояний ДО, поведение которого сопровождается многомерным временным рядом; сглаживания временного ряда; идентификации тренда фрагмента временного ряда;

- модель реконструкции состояний ДО на основе обобщенных эталонов и параметризации модели авторегрессии;

- модель выявления закономерностей (признаков и их весовых коэффициентов) на основе дискретного подхода к распознаванию состояний динамического объекта и формализма мультимножеств;

- модель выявления закономерностей как системы градаций квантования на основе максимального правдоподобия и теоретико-информационного подхода к распознаванию состояний;

- модель ранжирования динамических наборов альтернатив с целью корректного многокритериального оценивания методов и алгоритмов распознавания состояний ДО на основе нелинейной свертки критериев.

На основе перечисленных моделей предложены подходы к решению четырех прикладных задач: 1) построение модели наблюдателя при адаптивном управлении (АКАР) электро-механическим объектом; 2) экологическое диагностирование и прогнозирование; 3) моделирование состояния здоровья пациента для оценки и прогноза его состояния; 4) предсказание неизвестных значений атрибутов в базах данных.

Изложенные модели и соответствующие вычислительно реализуемые методы и алгоритмы могут служить основой автоматизированной обработки нестационарных временных рядов с нелинейным трендом.

Благодарности. Теоретические исследования и практическая апробация методов и реализующих их алгоритмов осуществлялись в соответствии с проектами РФФИ (№ 09-01-99014 р-офи, № 10-01-00462-а).

Большинство результатов книги докладывались на семинарах, руководимых д.т.н., профессором ТПУ В.Г. Букреевым, поставившим прикладную задачу

распознавания состояний технического объекта в реальном времени с целью применения апостериорной информации для коррекции управления⁴³.

Авторы выражают глубокую признательность В.В. Поддубному, доктору технических наук, профессору ТГУ, за внимательное прочтение рукописи и ценные и деликатные замечания, способствующие ее улучшению.

Авторы выражают благодарность Ю.Р. Цюю, к.т.н., доценту ТПУ, за выполнение расчетов (гл. 2, 5); А.Н. Мертвцову, аспиранту ТПУ, за разработку программного комплекса IReDSS (п. 8.3) в рамках работ по гранту РФФИ № 09-01-99014 р-офи; В.С. Лаходынову, аспиранту ТПУ, Г.В. Родионову, магистру ТПУ, за разработку 2-х моделей асинхронного двигателя в среде MatLab; И.С. Марочкиной, С.М. Бурковой, Г.Ю. Васильченко, Л.А. Казаковой и др. магистрам и бакалаврам кафедры ЭМИС ТУСУР, за создание программного обеспечения на основе алгоритмов глав 2, 9.

⁴³ Свидетельство на полезную модель «Система управления с распознаванием образов динамических состояний стохастического объекта» [144]

ЛИТЕРАТУРА

1. Anishchenko V.S., Pavlov A.N. Global reconstruction in application to multichannel communication // *Phys.Rev. E*. 1998. – V. 57. –P. 2455–2457.
2. Astolfi A., Ortega R. Immersion and invariance: A new tool for stabilization and adaptive control of nonlinear systems // *IEEE Trans. on Automatic Control*. – 2003. – Vol. 48. – No. 4. – P. 590–605.
3. Baake E., Baake M., Bock H.G., and Briggs K.M. Fitting ordinary differential equations to chaotic data // *Phys. Rev. A*, 1992. - V. 45. - No. 8. - P. 5524–5529.
4. Feigin A.M., Konovalov I.B., Molkov Y.I. Toward an understanding of the nonlinear nature of atmospheric photochemistry: essential dynamic model of the mesospheric photochemical system. // *J. Geophys. Res.* 1998. – V. 103. – № D19. – P. 25447–25460.
5. Granger C.W. Developments in the Study of Cointegrated Variables. - *Oxford Bulletin of Economics and Statistics*. - Vol. 48. - 1986. - pp. 213-228.
6. Györfi L., Kohler M. and Walk H. Weak and strong universal consistency of semirecursive kernel and partitioning regression estimates // *Statist. Decisions*. – 1998. – V. 16. – P. 1–18.
7. Ho D.L., Kolomeiseva M.B. Adaptive fuzzy logic control of robot-manipulator // *IFAC. Workshop on Manufacturing, Modeling, Management and Control*. – Prague, 2001. – P. 157–160.
8. Huang, et al The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis // *Proc. R. Soc. Lond. A*. – 1998. – T. 454. – C. 903–995.
9. Kira K., Rendell L. The Feature Selection Problem: Traditional Methods and a New Algorithm // *Proc. 10th Nat’l Conf. Artificial Intelligence (AAAI–92)*. – 1992. – P. 129–134.
10. Kolesnikova S.I. Estimation of Unknown Values of Continuous Attributes in Databases with Noise of Two Kinds/ *Pattern Recognition and Image Analysis: New Information Technologies (PRIA–9–2008)*. Proceedings of the 9th International Conference. Vol. 1. – Nizhni Novgorod, 2008. – P. 309–312.
11. Kolesnikova S.I., Yankovskaya A.E. Estimation of Significance of Features for Tests in Intelligent Systems// *Journal of Computer and Systems Sciences International*. – 2008. – Vol. 47, No 6. – 930–943. (ISSN 1064–2307, Pleiades Publishing, Ltd.)
12. Lorenz H.W. *Nonlinear Dynamical Economics and Chaotic. – Motion*. Berlin: Springer-Verlag. – 1993. – P. 319.
13. N.G. Zagoruiko, I.A. Borisova, V.V. Dyubanov, and O.A. Kutnenko. Methods of Recognition Based on the Function of Rival Similarity // *Pattern Recognition and Image Analysis*, 2008. – Vol. 18. – No. 1. – P. 1–6.

14. Narendra K.S., Annaswamy A. M. Robust adaptive control in the presence of bounded disturbances // *IEEE Trans. on Autom. Contr.* – 1986. – Vol. 31. – № 4. – P. 306–315.
15. Rulkov N.F. Regularization of synchronized chaotic bursts // *Phys. Rev. Lett.* 2001. – V. 86. – P. 183–186.
16. S.I. Kolesnikova. An integrated approach to evaluating the mutual influence of attributes in test recognition // *Cybernetics and Systems Analysis*. Springer New York. – 2009. – V. 45. – № 3. – С. 446–454.
17. Sauer T., Yorke J., Casdgli M. Embedology // *Journal of Statistical Physics*. – 1991. – Vol. 65. – № 3–4. – P. 579–616.
18. Takens F. Detecting strange attractors in turbulence // In: *Dynamical Systems and Turbulence. Lecture Notes in Mathematics*. – Berlin: Springer-Verlag, 1980. – V. 898. – P. 366–381.
19. W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. *Numerical Recipes*. - Cambridge, Cambridge University Press, 1997. – 1724 с.
20. Yankovskaya A., Kolesnikova S. An Approach to Calculation of Feature Weight Coefficients on the Base of Multisets Formalism in Intelligent Systems // *Knowledge-Based Software Engineering: proceedings of the Sixth Joint Conference*, IOS Press, 2004. – Vol. 108. – P. 159–168.
21. Yankovskaya A., Kolesnikova S. On Analysis of tests in Pattern Recognition Problem // *7th International Conference on Pattern Recognition and Image Analysis: New Information Technologies (PRIA-7-2004): conference proceedings*. – Vol. I. – St. Petersburg: SPbETU, 2004. – P. 141–144.
22. Yankovskaya A.E. Minimization of Orthogonal Disjunctive Normal Forms of Boolean Function to be Used as a Basis for Similarity and Difference Coefficients in Pattern Recognition Problems // *Pattern Recognition and Image Analysis*. – 1996. – V. 6. – No 1. – P. 60–61.
23. Yankovskaya A.E. Test Pattern Recognition with the Use of Genetic Algorithms // *Pattern Recognition and Image Analysis*. – 1999. – V. 9. – No. 1. – P. 121–123.
24. Yankovskaya A.E., Gedike A.I., Ametov R.V., Bleikher A.M. IMSLOG-2002 Software Tool for Supporting Information Technologies of Test Pattern Recognition // *Pattern Recognition and Image Analysis*. – 2003. – Vol. 13. – No. 4. – P. 650–657.
25. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. *Прикладная статистика: Исследование зависимостей*. – М.: Финансы и статистика, 1985. – 487 с.
26. Аносов О.Л., Бутковский О.Я., Кравцов Ю.А. Восстановление динамических систем по хаотическим временным рядам (краткий обзор) // *Изв. вузов. Прикладная нелинейная динамика*. – 2000. – Т. 8. – № 1. – С. 29–51.

27. Анохин А.М., Глотов В.А., Павельев В.В., Черкашин А.М. Методы определения важности критериев // Автоматика и телемеханика. – 1997. – № 8. – С. 3–31.
28. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование. – М.: Финансы и статистика, 2001. – 228 с.
29. Безручко Б.П., Смирнов Д.А. Математическое моделирование и хаотические временные ряды. – Саратов: ГосУНЦ «Колледж», 2005. – 320 с.
30. Безручко Б.П., Смирнов Д.А. Современные проблемы моделирования по временным рядам // Известия Саратовского госуниверситета, сер. Физика. – 2005. – Т.5. – Вып. 2.
31. Бериша А.М. Вагин В.Н., Куликов А.В., Фомина М.В. Методы обнаружения знаний в «зашумленных» базах данных // Искусственный интеллект. – 2005. – С. 143–158.
32. Бидюк П.И., Гасанов А.С., Шефер А.В., Зуев С.А. Алгоритм идентификации стохастических нелинейных систем по измерительным данным // Управл. системы и машины. – 2004. – №1. – С.12–18.
33. Бокс Дж., Дженкинс Г. Анализ временных рядов. Прогноз и управление. Вып.1. – М.: «Мир», 1971. – 406 с.
34. Болдин М.В., Симонова Г.И., Тюрин Ю.Н. Знаковый статистический анализ линейных моделей М.: Наука, Физматлит, 1997. – 288 с.
35. Борисов Ю.Ю. Построение прогнозирующих моделей динамических систем на основе исследования окрестностей реконструированных аттракторов // Автоматизация и современные технологии. – 2007. – № 2. – С. 32–37.
36. Букреев В.Г., Колесникова С.И., Цой Ю.Р. Информационная и программная поддержка распознавания состояний динамического объекта в условиях неопределенности // Интегрированные модели и мягкие вычисления в искусственном интеллекте: сб. научных трудов V-й Междунар. научно-практ. конф. – М.: Физматлит, 2009. – Т. 2.– С. 720–727.
37. Бутенин Н.В., Неймарк Ю.И., Фуфаев Н.А. Введение в теорию нелинейных колебаний. – М.: Наука, 1987. – 207 с.
38. В.В. Круглов, М.И. Дли, Р.Ю. Голупов. Нечеткая логика и искусственные нейронные сети. – М.: Физматлит, 2000. – 224 с.
39. Вагин В.Н. Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. – М.: Физматлит, 2004. – 704 с.
40. Вагин В.Н. Знание в интеллектуальных системах // Новости ИИ. 2002, №6. – С. 8–18.
41. Вапник В. Н., Червоненкис А. Я. Теория распознавания образов. – М.: Наука, 1974. – 416 с.

42. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. – 448 с.
43. Васильев В.А., Добровидов А.В., Кошкин Г.М. Непараметрическое оценивание функционалов от распределений стационарных последовательностей. – М.: Наука, 2004. – 508 с.
44. Васин Е.А., Костенко В.А., Коваленко Д.С. Автоматическое построение алгоритмов, основанных на алгебраическом подходе, для распознавания предаварийных ситуаций динамических систем // Искусственный интеллект. – 2006. – № 2. – С. 130–134.
45. Вильсон А.Дж. Энтропийные методы моделирования сложных систем. – М.: Наука, 1978. – 248 с.
46. Волченко Е.В. Модифицированный метод потенциальных функций // Бионика интеллекта – 2006. – № 1(64). – С. 86–92.
47. Воронцов К.В. Комбинаторная теория надёжности обучения по прецедентам: автореф. дис. докт. физ.-мат. наук. – Москва, 2010. – 42 с.
48. Галлагер Р. Теория информации и надёжная связь. – М., «Советское радио», 1974. – 720 с.
49. Генкин А.А. Новая информационная технология анализа медицинских данных. – СПб: Политехника, 1999. – 191 с.
50. Глухов В.В. Техническое диагностирование динамических систем – М.: Транспорт, 2000. – 96 с.
51. Грумондз В.Т. Динамика нелинейных систем: Некоторые задачи устойчивости и колебаний. – Вуз. книга, 2009. – 182 с.
52. Дюкова Е.В., Песков Н.В. Построение распознающих процедур на базе элементарных классификаторов // www.ccas.ru/frc/papers/djukova05construction.pdf.
53. Евланов Л.Г. Контроль динамических систем. – М.: Наука, 1979. – 431 с.
54. Журавлёв Ю.И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. – 1978. – Т. 33. – С. 5–68.
55. Журавлев Ю.И., Гуревич И.Б. Распознавание образов и анализ изображений // Искусственный интеллект в 3-х кн. Кн 2. Модели и методы: Справочник / Под ред. Д.А.Поспелова. – М: Радио и связь. – 1990. – С. 149–190.
56. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. – Москва: Фазис. – 2006. – 176 с.
57. Журбенко И.Г., Кожевникова И.А. Стохастическое моделирование процессов. – М.: МГУ, 1990. – 147 с.

58. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Изд. ИМ СО РАН, Новосибирск, 1999 г. – 273 с.
59. Зельдович Я.Б., Мышкис А.Д. Элементы прикладной математики. – М.: Главная редакция физико-математической литературы изд-ва «Наука», 1972. – 592 с.
60. Ивахненко А.А., Каневский Д.Ю., Рудева А. В., Стрижов В.В. Выявление групп объектов, описанных набором многомерных временных рядов // Математические методы распознавания образов (ММРО-13): сб. докладов 13-й Всеросс. конф. – М.: МАКС Пресс, 2007. – С. 134–137.
61. Изосимов Д.Б., Рывкин С.Е. Идентификация частоты вращения и составляющих вектора потокосцепления ротора асинхронного двигателя по измерениям токов и напряжений обмоток статора // Электричество. – 2005. – № 4. – С. 32–40.
62. Инишева Л.И., Пяткова (Колесникова) С.И., Дементьева Т.В.. Автоматизированная информационная система «Химия торфов» (АИС «Химия торфов»). Свидетельство об официальной регистрации базы данных №970002, 05.01.1997.
63. Инишева Л.И., Пяткова (Колесникова) С.И., Савичева О.Г. Автоматизированная информационная система «Торфяные ресурсы» (АИС «Торфяные ресурсы»). Свидетельство об официальной регистрации базы данных №970005, 20.01.1997.
64. Кашьяп Р.Л., Рао А.Р. Построение динамических стохастических моделей по экспериментальным данным. – М.: Наука, 1985. – 384 с.
65. Кедрин В.С., Сальникова М.К. Сравнительный анализ методов спектрального и сингулярного разложения в задачах прогнозирования состояния сложных динамических систем // Труды Братского государственного университета, БрГУ, 2007. – Т. 2. – С. 45–49.
66. Кендалл Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1986. – 550 с.
67. Киричков В.Н. Идентификация объектов систем управления технологическими процессами / В.Н. Киричков. Киев: Высшая школа, 1990. – 218 с.
68. Клир Дж. Системология. Автоматизация решения системных задач: Пер. с англ. / Дж. Клир; пер. М.А. Зуев, ред. пер А.И. Горлин. – М. : Радио и связь, 1990. – 538 с.
69. Ключев В. И. Теория электропривода: учеб. для вузов. – 3-е изд. перераб. и доп. – М.: Энергоатомиздат, 2001. – 704 с.
70. Колесников А.А. Основы синергетической теории управления. – М.: ИСПО-Сервис, 2000. – 260 с.
71. Колесникова С.И. Выявление закономерностей во временных рядах при распознавании состояний сложных объектов управления // Приборы и системы. Управление, контроль, диагностика. – 2010. – № 5.

72. Колесникова С.И. Метод парных сравнений при оценивании динамических наборов альтернатив // Качество. Инновации. Образование. – 2010. – № 4.

73. Колесникова С.И. Методы анализа информативности разнотипных признаков // Вестник Томского гос. ун-та. Управление, вычислительная техника и информатика. – 2009. – № 1(6). – С. 69–80.

74. Колесникова С.И. Методы оценивания зависимых признаков при классификации многопризнаковых объектов // Вестник Томского гос. ун-та. Приложение. – 2007. – № 23. – С. 315–320.

75. Колесникова С.И. Методы распознавания состояний динамических систем // Известия ТПУ. – 2010. – Т. 316. – № 5. – С. 55–62.

76. Колесникова С.И. Модификация метода анализа иерархий для динамических наборов альтернатив // Прикладная дискретная математика. – 2009. – № 4(6). – С. 102–109.

77. Колесникова С.И. О подходах к оцениванию информативности признаков в тестовом распознавании // Известия ТПУ. – 2006. – № 8. – С. 23–28.

78. Колесникова С.И. Особенности применения линейной свертки критериев в методе парных сравнений // Информационные технологии. – 2010. – № 12.

79. Колесникова С.И. Системный подход к оцениванию взаимного влияния признаков в тестовом распознавании // Кибернетика и системный анализ. – 2009. – № 3. – С. 127–135.

80. Колесникова С.И., Букреев В.Г. Информационный подход к распознаванию состояний динамической системы // Кибернетика и высокие технологии XXI века (С&Т-2009): сб. докладов X Междунар. научно-практ. конф. – Воронеж: НПФ «Саквоее» ООО, 2009. – Т. 1. – С. 54–64.

81. Колесникова С.И., Букреев В.Г. Подход к распознаванию состояний технической системы на основе энтропии и формализма мультимножеств // Прикладная дискретная математика. Приложение. – 2009. – № 1. – С. 109–111.

82. Колесникова С.И., Букреев В.Г. Распознавание состояний динамической системы // Цифровая обработка сигналов и ее применение: труды Росс.научно-технич.общества радиотехники, электроники и связи им. А.С.Попова, доклады 11-й Междунар.конф. Выпуск: XI – 2. – 2009. – С. 619–622.

83. Колесникова С.И., Букреев В.Г., Мертвецов А.Н., Цой Ю.Р. Информационная система для распознавания состояний стохастической системы // Программные продукты и системы. – 2010. – № 4.

84. Колесникова С.И., Букреев В.Г., Цой Ю.Р., Мертвецов А.Н. Информационно-диагностическое программное средство для распознавания состояний зашумленного динамического объекта // Информационные технологии в науке, образовании, телекоммуникации, бизнесе (IT+S&E`09): материалы XXXVI Ме-

ждународ. конф. – Украина, Крым, Ялта–Гурзуф: Приложение к журналу «Открытое образование», 2009. – С. 92–94.

85. Колесникова С.И., Волченко Е.В. Подход к сглаживанию и реконструкции формы состояний стохастического объекта на основе модели эталонов // Известия ТПУ. – 2010. – Т. 316. – № 5. – С. 34–40.

86. Колесникова С.И., Лаходынов В.С., Цой Ю.Р. Исследование качества распознавания состояний стохастической системы // Информационные технологии. – 2010. – № 6. – С. 56–62.

87. Колесникова С.И., Мертвецов А.Н., Цой Ю.Р. Информационно-аналитическая поддержка обработки зашумленных временных рядов // Компьютерные науки и технологии: сб. трудов Первой Междунар. научно-практ. конф. – Белгород: ГиК, 2009. – Ч. 1:– С. 186–190.

88. Колесникова С.И., Можейко В.И., Цой Ю.Р., Янковская А.Е. Алгоритмы выбора оптимального множества безызбыточных диагностических тестов в интеллектуальных системах поддержки принятия решений // Системный анализ и информационные технологии (САИТ-2005): труды первой Междунар. конф. в 2 т. – М.: КомКнига, – 2005. Т. 1. – С. 256–262.

89. Колесникова С.И., Цой Ю.Р. Оценка качества распознавания состояний динамического объекта // Математические методы распознавания образов (ММРО-14): сб. докладов 14-й Всеросс. конф. – М.: МАКС Пресс, 2009. – С. 368–371.

90. Колесникова С.И., Шапцев В.А. Подход к решению задачи оценивания состояний экологического объекта // Экологические системы и приборы. – 2010. – № 8.

91. Колесникова С.И., Янковская А.Е. Выбор оптимального подмножества тестов с применением редукции многокритериального выбора и метода анализа иерархий // Системный анализ и информационные технологии (САИТ-2009): труды третьей Междунар. конф. – М.: URSS, 2009. – С. 278–284.

92. Колесникова С.И., Янковская А.Е. К вопросу вычисления весовых коэффициентов признаков в интеллектуальных системах поддержки принятия решений при большой размерности признакового пространства // Вестник Томского государственного университета. Приложение. – 2006. – № 18 – С. 223–227.

93. Колесникова С.И., Янковская А.Е. К вычислению весовых коэффициентов признаков в интеллектуальных динамических системах // 10-я национальная конференция по искусственному интеллекту с международным участием (КИИ-06). – Москва, Физматлит, – 2006. – Т. 3.– С. 783–791.

94. Колесникова С.И., Янковская А.Е. К определению весовых коэффициентов признаков в системах тестового распознавания // Интеллектуальные системы (AIS'06), Интеллектуальные САПР (CAD-2006): труды Междунар. научно-практ. конф. – Москва: Физматлит, 2006. – Т. 1. – С. 335–344.

95. Колесникова С.И., Янковская А.Е. Методы оценивания весовых ко-

эффицентов признаков в интеллектуальных системах поддержки принятия решений // IX Национальная конференция по искусственному интеллекту с международным участием (КИИ-2004): сб. научных трудов. – Москва: Физматлит, 2004. – Том 2. – С. 833–841.

96. Колесникова С.И., Янковская А.Е. Основанные на мультимножествах методы определения весовых коэффициентов признаков в интеллектуальных системах // Интеллектуальные системы (AIS'07), Интеллектуальные САПР (CAD-2007): труды Междунар. научно-практ. конф. – Москва: Физматлит, 2007. – Т. 2. – С. 198–206.

97. Колесникова С.И., Янковская А.Е. Оценивание зависимых признаков в интеллектуальных системах // Системный анализ и информационные технологии (САИТ-2007): труды второй Междунар. конф. в 2 т. – Т. 1. – М.: URSS, 2007. – С. 137–140.

98. Колесникова С.И., Янковская А.Е. Оценивание признаков в интеллектуальных тестовых распознающих системах на основе энтропии // Интеллектуальный анализ информации (ИАИ-2007): сб. трудов VII Междунар. конф. – Киев: «Просвіта», 2007. – С. 159–168.

99. Колесникова С.И., Янковская А.Е. Оценка значимости признаков для тестов в интеллектуальных системах // Известия РАН. Теория и системы управления. – 2008. – № 6. – С. 135–148.

100. Колесникова С.И., Янковская А.Е. Статистический подход к оцениванию зависимых признаков в интеллектуальных системах // Математические методы распознавания образов (ММРО-13): сб. докладов 13-й Всеросс. конф. – М.: МАКС Пресс, 2007. – С. 143–146.

101. Колмогоров А.Н. К вопросу о пригодности найденных статистическим путем формул прогноза. – М.: Заводская лаборатория, 1933. – № 1. – С. 164–167.

102. Коломейцева М.Б., Хо Д.Л. Адаптивные системы управления динамическими объектами на базе нечётких регуляторов. М.: Спутник+, 2002. – 140 с.

103. Копылов И.П. Математическое моделирование электрических машин. – М.: Высш. шк. – 2001. – 327 с.

104. Кошкин Г.М., Лаходынов В.С. Полурекуррентная непараметрическая идентификация условных функционалов слабозависимых последовательностей // Вестник ТГУ. Управление, вычислительная техника и информатика, № 1(2), 2008. – С. 57–68.

105. Красовский А.А. Математическая и прикладная теория. Избранные труды. – М.: Наука, 2002. – 362 с.

106. Красовский А.А. Современная прикладная теория управления (в 3-х частях). Ч. I. «Оптимизационный подход в теории управления» / Под ред. А.А. Колесникова. – Таганрог: ТРТУ, 2000. – С. 268–311.

107. Кремер Н.Ш., Путко Б.А. Эконометрика. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.
108. Кудрявцев В.Б., Андреев А.Е. Теория тестового распознавания. – Интеллектуальные системы. – 2006. – Т. 10. – вып.1–4. – С. 95-140.
109. Кузнецов П.И., Пчелинцев Л.А., Гайденок В.С. Контроль и поиск неисправностей в сложных системах. – Советское Радио, 1969. – 240 с
110. Кук Ю.В., Лаврикова Е.И. Спектральный метод распознавания состояний динамических систем // Комп'ютерні засоби, мережі та системи. – 2007. – № 6 – С. 133–140.
111. Кулебакин В.С. О поведении непрерывно возмущаемых автоматизированных линейных систем // Доклады АН СССР. – 1949. – Т. 68. – № 5. – С. 73–79.
112. Кульбак С. Теория информации и статистика. – М.: – Наука, 1967. – 408 с.
113. Ларичев О.И. Теория и методы принятия решений, а также Хроника событий в Волшебных странах: Учебник. Изд. третье, перераб. и доп. – М.: Университетская книга, Логос. – 2006. – 296 с.
114. Лачинов В.М., Поляков А.О. Информодинамика или путь к миру открытых систем. – СПб.: Изд-во СПбГТУ, – 1999. – 196 с.
115. Луценко Е.В. Разработка методологии синтеза адаптивных АСУ сложными объектами на основе применения моделей распознавания образов и принятия решений. Дис... к.т.н. – Краснодар: КубГТУ. 1999. – 187 с.
116. Льюнг Л. Идентификация систем. Теория для пользователей. – М.: Наука, 1991. – 432 с.
117. Мирошник И.В., Никифоров В.О., Фрадков А.Л.. Нелинейное и адаптивное управление сложными динамическими системами. – СПб.: Наука, 2000. – 562 с.
118. Михайлов Ю.Б. Математические основы повышения точности прогнозирования количественных характеристик процессов (в технике, экономике, экологии, социологии, бизнесе). – М.: ООО Издательство «Научтехлитиздат», 2000. – 206 с.
119. Моисеев Н.Н. Математические задачи системного анализа. – М.: Наука, 1981. – 488 с
120. Моттль В.В., Мучник И.Б. Скрытые марковские модели в структурном анализе сигналов. – М.: Физматлит, 1999. – 352 с.
121. Неймарк Ю. И., Теклина Л. Г. Анализ фазовых траекторий многомерных динамических систем методами распознавания на основе одномерных временных рядов // Математические методы распознавания образов (ММРО–13): доклады XIII Всеросс. конф. Сборник докладов. – М.: МАКС Пресс, 2007. – С. 191 – 193.

122. Нестеров И.И., Кожевников Ю.А., Рядинский В.Ю. Задачи искусственного интеллекта в экологическом прогнозировании // Доклады академии наук. - 2007. – Т. 414. - № 15. – С. 52-56.

123. Новиков А.А. О времени выхода сумм ограниченных случайных величин из криволинейной полосы // Теория вероятностей и ее применения. – 1981. – Т. 26. – № 2. – С. 287–301.

124. Новиков А.А. Об оценках и асимптотическом поведении вероятностей непересечения подвижных границ суммами независимых случайных величин // Изв. АН СССР. Сер. Математика. – 1980. – Т. 44. – № 4. – С. 868–885.

125. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат, 1991. – 304 с.

126. Ногин В.Д. Упрощенный вариант метода анализа иерархий на основе нелинейной свертки критериев // Журнал вычислительной математики и математической физики. – 2004. – Т.44. – № 7. – С. 1259-1268.

127. Орлов Ю.Н., Осминин К.П. Анализ нестационарных временных рядов / Препринт ИПМ им. М.В. Келдыша РАН. – 2007. – № 36. – 24 с.

128. Орлов Ю.Н., Осминин К.П. Построение выборочной функции распределения для прогнозирования нестационарного временного ряда // Математические модели, 2008. – № 9. – С. 23-33.

129. Пащенко Ф.Ф. Введение в состоятельные методы моделирования систем. В 2-х ч. Ч. 2. Идентификация нелинейных систем. – М.: Финансы и статистика, 2007. – 288 с.

130. Перегудов Ф.И., Тарасенко Ф.П. Введение в системный анализ. – М.: Высшая школа, 1989. – 367 с.

131. Петровский А.Б. Упорядочивание и классификация объектов с противоречивыми признаками // Новости искусственного интеллекта. – 2003. – № 4. – С. 34–43.

132. Подиновский В.В. Количественная важность критериев // Автоматика и телемеханика. – 2000. – № 5. – С. 110–123.

133. Подиновский В.В., Ногин В.Д. Парето-оптимальные решения многокритериальных задач. – Н.: Наука. – 2007. – 255 с.

134. Поляк Б.Т., Щербаков П.С. Робастная устойчивость и управление. М.: Наука, 2002. – 303 с.

135. Попков Ю.С., Киселев О.Н., Петров Н.П. Идентификация и оптимизация нелинейных стохастических систем. – М Энергия, 1976. – 440 с.

136. Поспелов Д.А. Данные и знания. Искусственный интеллект: в 3 кн. Кн. 2. Модели и методы: Справ. – М.: Радио и связь, 1990. – 304 с.

137. Прангишвили И.В. Энтропийные и другие системные закономерности: Вопросы управления сложными системами. – М.: Наука, 2003. – 428 с.

138. Пяткова (Колесникова) С.И. О двусторонней оценке надежности технических систем системы с программным резервированием // Надежность и контроль качества. – 1991. – №3.- С. 9-14.
139. Рудаков К.В., Чехович Ю.В. Алгебраический подход к проблеме синтеза обучаемых алгоритмов выделения трендов // Доклады РАН. - 2003. -Т. 388. - № 1.- С. 33-36.
140. С.И. Колесникова, А.Е. Янковская. Вероятностные подходы к оцениванию информативности признаков в интеллектуальных системах принятия решений Междунар // Труды Всероссийской ФАМ' 2008 конф. – Красноярск, Сиб. Федер. Университет. – 2008. – ч. 2. – С. 109–115.
141. Саати Т.Л. Принятие решений. Метод анализа иерархий. – М.:Радио и связь. – 1989. – 311 с.
142. Самохвалов Ю.Я. Групповой учет относительного превосходства альтернатив в задачах принятия решений // Кибернетика и системный анализ. – 2003. – № 6. – С. 141–145.
143. Самохвалов Ю.Я. Особенности применения метода анализа иерархий при оценке проблем по метрическим критериям // Кибернетика и системный анализ. – 2004. – № 5. – С. 15–19.
144. Свидетельство на полезную модель 86021 РФ, № 2009114963/22, МПК G05B13/00. Система управления с распознаванием образов динамических состояний стохастического объекта / В.Г. Букреев, С.И. Колесникова. Заявлено 20.04.09. Опубл. 20.08.09. Бюл. № 23. – 3 с.
145. Свидетельство об официальной регистрации программы ЭВМ № 2010620277. База данных результатов электромагнитных и акустических измерений геофизических процессов / А.А. Беспалько, В.Г. Букреев, С.И. Колесникова, А.Н. Мертвецов, Л.В. Яворович. Заявка 2010620117. Дата поступления. 19.03.10. Зарегистр. 14.05.2010.
146. Свидетельство об официальной регистрации программы ЭВМ №2010610441. Интеллектуальная модель распознавания состояний динамических систем «IReDDS» / С.И. Колесникова, В.Г. Букреев, А.Н. Мертвецов Ю.Р. Цой, В.С. Лаходынов. Заявка 2009616425. Дата поступл. 16.11.09. Зарегистр. 11.01.2010.
147. Семёнычев В.К. Идентификация экономической динамики на основе моделей авторегрессии. – Самара.:АНО «Изд. СНЦ РАН». – 2004. – 243 с.
148. Симанков В.С., Луценко Е.В. Синтез адаптивных АСУ сложными системами с применением моделей распознавания образов // <http://lc.narod.ru/aidos/Public/St60.htm>
149. Синергетика и проблемы теории управления: сборник научных трудов / Под ред. А.А. Колесникова. – М.: ФИЗМАТЛИТ, 2004. – 504 с.

150. Стрижов В.В. Методы индуктивного порождения регрессионных моделей. – М.: ВЦ РАН, 2008. – 62 с.
151. Татт У.Т. Теория графов. - Москва: Мир, 1988. – 314 с.
152. Темников Ф.Е., Афонин В.А., Дмитриев В.И. Технические основы информационной техники. – М.: Энергия, 1979. – 511 с.
153. Терехов С.А. Нейроинформатика / А. Н. Горбань, В. Л. Дунин-Барковский, А. Н. Кирдин и др. – Новосибирск: Наука. Сибирское предприятие РАН, 1998. – 296 с.
154. Терпугов А.Ф., Колупаева С.Н., Янковская А.Е. Оценка информативности признаков и объема обучающей выборки в задаче распознавания образов // Математические методы распознавания образов (ММРО–9): доклады 11-й Всероссийской конференции. – Москва, 1999. – С. 108–111.
155. Тимашев С.А. Управление рисками взаимосвязанных систем критичных инфраструктур и территорий / Безопасность критичных инфраструктур и территорий: тр. II Всероссийской науч.-техн. конф. – Екатеринбург: УрО РАН, 2008. – С. 7–9.
156. Ту Дж., Гонсалес Р. Принципы распознавания образов / Пер. с англ. Под ред. Ю.И. Журавлева. – М.: Мир, 1978. – 411 с.
157. Тырсин А.Н. Идентификация нестационарных экономических процессов на основе дискретно-совпадающих моделей авторегрессии // Известия Уральского государственного экономического университета. – 2004. – № 9. – С. 44–51.
158. Тюкин И.Ю., Терехов В.А. Адаптация в нелинейных динамических системах / Санкт-Петербург: ЛКИ, 2008. – 384 с.
159. Финк Л.М. Сигналы, помехи, ошибки. – М.: Радио и Связь, 1984. – 256 с.
160. Флеминг У., Ришел Р. Оптимальное управление детерминированными и стохастическими системами. – М.: Наука, 1978. – 316 с.
161. Хайкин С. Нейронные сети. Полный курс (пер. с англ.) – М.: «Вильямс», 2006. – 1104 с.
162. Шахнов И.Ф. Количественная оценка важности целей // Известия РАН. Теория и системы управления. – 2003. – № 1. – С. 78–86.
163. Ширяев А.Н. Вероятность. – М.: Наука, 1980. – 575 с.
164. Шоломов Л.А. Исследование одного класса динамических процедур коллективного выбора // Нелинейная динамика и управление. Вып. 5. – М.: Физматлит, 2007. – 400 с.
165. Шоломов Л.А. О собственной информации нечетких текстов // Нелинейная динамика и управление. Вып. 6. – М.: Физматлит, 2008. – 340 с.
166. Шугай Ю.С. Нейросетевые алгоритмы прогнозирования событий и поиска предвестников в многомерных временных рядах // Искусственный интеллект. – Донецк, 2004. – № 2. – С. 211-215.

167. Шурыгин А.М.. Прикладная стохастика. - М.:Финансы и статистика, 2005. – 224 с.
168. Шустер Г. Детерминированный хаос. Введение. – М.:Мир, 1988. – 240 с.
169. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. – М.: Финансы и статистика, 1988. – 263 с.
170. Яковлев Г.Н. Функциональные пространства. – М.: МФТИ, 2000. – 128 с.
171. Янковская А.Е. Колесникова С.И. Сравнительный анализ методов получения весовых коэффициентов признаков в задаче распознавания образов // Математические методы распознавания образов (ММРО–11): доклады 11-й Всероссийской конференции. – Москва, 2003. – С. 227–229.
172. Янковская А.Е., Колесникова С.И. Методы определения весовых коэффициентов признаков в интеллектуальных системах // Вестник ТГУ. Приложение. – 2004. – № 9 (II). – С. 76–83.
173. Янковская А.Е., Колесникова С.И. О применении мультимножеств к задаче вычисления весовых коэффициентов признаков в интеллектуальных распознающих системах // Искусственный интеллект. Украина, Донецк: Наука і освіта. – 2004. – № 2. – С. 216–220.
174. Янковская А.Е., Колесникова С.И. Оценивание результативности трудовой деятельности работников федеральных бюджетных учреждений // Успехи современного естествознания. – 2008. – № 9. – С. 39–41.
175. Янковская А.Е. Логические тесты и средства когнитивной графики в интеллектуальной системе // Новые информационные технологии в исследовании дискретных структур: доклады 3-ей Всерос. конф. с междунар. участием. – Томск: Изд-во СО РАН, 2000. – С. 163–168.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ к главе 2

Доказательство утверждения 2.1.

Представим Z-преобразование функции $x_k=f(k\Delta)$ дробно-рациональной функцией:

$$X(z) = \sum_{k=0}^{\infty} x_k z^{-k} = \frac{P(z)}{Q(z)} = \frac{\sum_{j=0}^{p'} p_j z^{-j}}{\sum_{j=0}^{q'} q_j z^{-j}}, p' \leq q' < \infty. \quad (\text{П2.1})$$

Преобразуем равенство (П2.1): $X(z) \frac{Q(z)}{P(z)} = 1$. На основании теоремы о делении многочленов (с остатком) для любых двух многочленов $Q(x)$ и $P(x)$ можно найти такие многочлены $G(x)$ и $R(x)$, что $Q(x)=G(x)P(x)+R(x)$, причем степень $R(x)$ меньше степени $G(x)$ или же $R(x)=0$, а многочлены $G(x)$ и $R(x)$ определяются однозначно. Отсюда следует справедливость представления:

$$X(z) \left(\sum_{j=0}^{p'} b_{1j} z^{-j} \sum_{j=0}^{p'} b_{2j} z^{-j} + \sum_{j=0}^{r'} r_j z^{-j} \right) = 1, \quad \text{или} \quad X(z) \left(\sum_{j=0}^p b_j z^{-j} + \sum_{j=0}^{r'} r_j z^{-j} \right) = 1,$$

$p > r'$. Раскрыв скобки и, применяя к полученному выражению обратное Z-преобразование, получим на основе свойства задержки Z-преобразования:

$$\sum_{j=0}^p a_j x(k-j) = \delta(k), \quad \text{где } \delta(k) \text{ – символ Кронекера. Отсюда следует справедли-$$

вость 1-й части утверждения 2.1: $x_k = \sum_{j=1}^p a_j x_{k-j}, \forall n \geq p$.

Справедливость 2-й части утверждения 2.1 следует из справедливости равенств для моментов наблюдений $k, \dots, k-p+1, k \geq p$:

$$x_k = y_k - \xi_k, \dots, x_{k-p+1} = y_{k-p+1} - \xi_{k-p+1},$$

подставив которые в $y_k = x_k + \xi_k, k \geq p$, получим параметрическую авторегрессию:

$$y_k = \sum_{j=1}^p a_j x_{k-j} + \xi_k = \sum_{j=1}^p a_j (y_{k-j} - \xi_{k-j}) + \xi_k = \sum_{j=1}^p a_j y_{k-j} + \xi_k - \sum_{j=1}^p a_j \xi_{k-j}.$$

Утверждение 2.1 доказано.

Доказательство утверждения 2.2.

В нормированном пространстве два элемента равны, если норма их разности равна нулю. Нетрудно проверить, что функция $r(x, y) = \|x-y\|, x, y \in X \subset L$, удовлетворяет всем аксиомам метрики, т.е. линейное нормированное пространство является метрическим, и поэтому в нём всегда можно упорядочить величины: $r(f, f_1), r(f, f_2), \dots, r(f, f_n)$, где $r(f, f_n) = \|f - f_n\|$:

$$r(f, f_{j1}) \leq r(f, f_{j2}) \leq \dots \leq r(f, f_{jn}).$$

Таким образом, выбор «лучшей» аппроксимирующей функции для f связан с нахождением функции f_{j1} , расстояние до которой минимально.

Утверждение 2.2 доказано.

Доказательство утверждения 2.3.

Необходимым условием минимальности функции $\sum_{l=1}^{\kappa_i} \phi \left(\left| y'_l - \sum_{j=1}^{p_i} a_j^i y'_{l-j} \right| \right)$

по искомым оценкам параметров является выражение:

$$\sum_{l=1}^{\kappa_i} \dot{\phi}(\lambda_l) \text{sign} \left(y'_l - \sum_{j=1}^{p_i} a_j^i y'_{l-j} \right) \left(- \sum_{j=1}^{p_i} y'_{l-j} \right) = 0, \quad \lambda_l = \left| y'_l - \sum_{j=1}^{p_i} a_j^i y'_{l-j} \right|.$$

Отсюда следует, что задача поиска оценок $(a_1^i, \dots, a_{p_i}^i)$ эквивалентна задаче поиска решений системы:

$$\sum_{j=1}^{p_i} \sum_{l=1}^{\kappa_i} \dot{\phi}(\lambda_l) y'_{l-j} \text{sign} \left(y'_l - \sum_{j=1}^{p_i} a_j^i y'_{l-j} \right) = 0. \quad (\text{П2.2})$$

Из работы [34] следует, что оценки $(a_1^i, \dots, a_{p_i}^i)$, получаемые решением системы (П2.2) при условии $\sup_{y \in R^p} |y y^T \varphi(y)| < \infty$:

$$\sum_{j=1}^p \sum_{l=1}^{\kappa} \varphi(y_{l-1}, \dots, y_{l-p}) y_{l-j} \text{sign} \left(y_l - \sum_{j=1}^p a_j y_{l-j} \right) = 0,$$

являются устойчивыми к большим ошибкам. Отсюда следует, что оценки $(a_1^i, \dots, a_{p_i}^i)$, получаемые решением системы (П2.2) при условии $\sup_{x \in R^+} |x^2 \phi'(x)| < \infty$, обладают таким же свойством.

Утверждение 2.3 доказано.

ПРИЛОЖЕНИЕ к главе 3

Доказательство утверждения 3.1.

Воспользуемся свойствами матрицы парных сравнений (МПС) b и c (3.1), из которых следует справедливость представления каждого элемента b_{ij} в виде:

$$b_{ij} = b_{i1} \cdot b_{1j} = b_{1j} (b_{1i})^{-1}.$$

Отсюда следует, что МПС $\mathbf{B} = \|b_{ij}\|$ представима в виде:

$$\mathbf{B} = \begin{bmatrix} 1 & b_{12} & \dots & b_{1g} \\ b_{21} & 1 & \dots & b_{2g} \\ \dots & \dots & \dots & \dots \\ b_{g1} & b_{g2} & \dots & 1 \end{bmatrix} = \begin{bmatrix} 1 & b_{12} & \dots & b_{1g} \\ 1/b_{12} & 1 & \dots & b_{1g}/b_{12} \\ \dots & \dots & \dots & \dots \\ 1/b_{1g} & b_{12}/b_{1g} & \dots & 1 \end{bmatrix} =$$

$$= \frac{1}{b_{12}} \frac{1}{b_{13}} \dots \frac{1}{b_{1g}} \begin{bmatrix} 1 & b_{12} & \dots & b_{1g} \\ 1 & b_{12} & \dots & b_{1g} \\ \dots & \dots & \dots & \dots \\ 1 & b_{12} & \dots & b_{1g} \end{bmatrix} = \frac{b_{12} b_{13} \dots b_{1g}}{b_{12} b_{13} \dots b_{1g}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}.$$

Далее доказательство утверждения непосредственно следует из вида характеристического уравнения для \mathbf{B} , имеющего место при выполнении этого свойства матрицы парных сравнений (3.1):

$$\det(\mathbf{B} - \lambda \mathbf{E}) = \det \begin{bmatrix} (1-\lambda) & 1 & \dots & 1 \\ 1 & (1-\lambda) & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & (1-\lambda) \end{bmatrix} = (-1)^g \lambda^{g-1} (\lambda - g) = 0,$$

из которого следует, что матрица \mathbf{B} будет иметь только два собственных значения (при выполнении условий (3.1)), при этом, значение g является максимальным. В противном случае собственный вектор несовместной МПС будет соответствовать максимальному собственному значению, которое строго больше g , а не равно g .

Доказательство теоремы 3.1.

Не ограничивая общности, рассмотрим в качестве множеств Θ_1 и Θ_2 , $\Theta_1 \subset \Theta_2$, следующие: $\Theta_1 = \{z_1, z_2, \dots, z_{g-1}\}$, $\Theta_2 = \{z_1, z_2, \dots, z_{g-1}, z_g\}$. Обозначим матрицы парных сравнений (МПС) признаков множества Θ_1 через $\mathbf{A}' = \|a'_{ij}\|_{(g-1) \times (g-1)}$, ($a'_{ij} = w'_i/w'_j$), и МПС признаков множества Θ_2 , $\mathbf{A} = \|a_{ij}\|_{g \times g}$, ($a_{ij} = w_i/w_j$), соответственно. В качестве приближения собственных векторов МПС \mathbf{A}' , \mathbf{A} в МАИ используют векторы $\mathbf{w}' = (w'_1, w'_2, \dots, w'_{g-1})^T$, $\mathbf{w} = (w_1, w_2, \dots, w_g)^T$, компоненты которого вычисляются как среднее геометрическое по элементам строк, т.е.

$$w'_i = \left(\prod_{l=1}^{g-1} a'_{il} \right)^{1/(g-1)} \text{ и } w_i = \left(\prod_{l=1}^g a_{il} \right)^{1/g}, \text{ соответственно.}$$

Пусть после применения МАИ оценки альтернатив таковы (ВКА), что $w_i > w_j$, то есть на множестве Θ_1 задано отношение предпочтения $z_i \succ z_j$. Покажем, что это отношение предпочтения ($z_i \succ z_j$) в общем случае не сохраняется на множестве Θ_2 .

Очевидно, что для сохранения отношения $z_i \succ z_j$ на множестве Θ_2 необходимо должно выполняться неравенство: $w_i > w_j$, т.е.

$$\prod_{l=1}^g a_{il} > \prod_{l=1}^g a_{jl}. \quad (\text{ПЗ.1})$$

Заметим, что $\prod_{l=1}^g a_{il} = \prod_{l=1}^{g-1} a'_{il} \cdot a_{ig}$, $\prod_{l=1}^g a_{jl} = \prod_{l=1}^{g-1} a'_{jl} \cdot a_{jg}$, $i, j < g$, тогда неравенство (ПЗ.1) примет вид: $D'_{i,g-1} \cdot a_{ig} > D'_{j,g-1} \cdot a_{jg}$, $i, j < g$, где $D'_{s,g-1} = \prod_{l=1}^{g-1} a'_{sl}$, $s \in \{i, j\}$. Отсюда следует, что для справедливости отношения (ПЗ.1) должно выполняться неравенство:

$$\frac{a_{ig}}{a_{jg}} > \frac{D'_{j,g-1}}{D'_{i,g-1}}, \quad i, j < g. \quad (\text{ПЗ.2})$$

Далее, следуя технике МАИ, используем свойства МПС (3.1) *a*)-*d*).

Преобразуем выражение (ПЗ.2), осуществив соответствующие замены, применяя свойства (3.1) *b*) и *c*):

$$\frac{a_{ig}}{a_{jg}} = \frac{a_{ij} \cdot a_{ig}}{a_{jg}} = a_{ij} = a'_{ij}; \quad \frac{D'_{j,g-1}}{D'_{i,g-1}} = \frac{\prod_{l=1}^{g-1} a'_{jl}}{\prod_{l=1}^{g-1} a'_{il}} = \prod_{l=1}^{g-1} \frac{a'_{jl}}{a'_{il}} = \prod_{l=1}^{g-1} \frac{a'_{ji} \cdot a'_{il}}{a'_{il}} = \prod_{l=1}^{g-1} \frac{1}{a'_{ij}} = (a'_{ij})^{1-g}.$$

Таким образом, для справедливости отношений (ПЗ.1) и (ПЗ.2) необходимо должно выполняться неравенство: $a_{ij} > (a_{ij})^{1-g}$, $i, j < g$, или: $(a_{ij})^g > 1$, $i, j < g$, что, очевидно, будет иметь место только в частных случаях.

Теорема 3.1 доказана.

Доказательство теоремы 3.2.

Доказательство аналогично доказательству теоремы 1. Полагаем для определенности $z_i \succ z_j$, для $z_i, z_j \in \Theta_1$, т.е. $w'_i > w'_j$, откуда следует, что при выполнении условий теоремы 2 и свойств МПС *a*)-*d*) выполняется неравенство $w_i/w_j > 1$, т.е. $w_i > w_j$, и $z_i \succ z_j$, для $z_i, z_j \in \Theta_2$. Теорема 3.2 доказана.

Доказательство теоремы 3.3.

При выполнении условия (1) теоремы 3.2 справедлива полезная формула, связывающая ВКА двух множеств (по каждому критерию): $w_i = w_i^{(g-1)/g} a_{ig}^{1/g}$, $i < g$. Действительно, использование представления компонент весового вектора альтернатив в МАИ как среднего геометрического по строкам МПС приводит к справедливости соотношений:

$$w_i = \left(\prod_{l=1}^g a_{il} \right)^{1/g} = \left(\prod_{l=1}^{g-1} a_{il} \right)^{1/g} a_{ig}^{1/g} = \left(\prod_{l=1}^{g-1} a'_{il} \right)^{1/g} a_{ig}^{1/g} = w_i^{(g-1)/g} a_{ig}^{1/g}.$$

Пусть для определенности $z_i \succ z_j$, для $z_i, z_j \in \Theta_1$, т.е. $w'_i > w'_j$. Требуется показать, что это отношение предпочтения $z_i \succ z_j$ сохранится на множестве Θ_2 ($i, j < g$) при оценивании весовых коэффициентов альтернатив по модифицированному МАИ, т.е. необходимо должно выполняться неравенство: $w_i > w_j$ ($i, j < g$).

Предположим сначала выполнение условия (1) теоремы 3.2. Применяя полученную зависимость, вышеприведенные формулы для ВКА по методу ММАИ, свойства *a)-d)* (3.1) МПС, получаем:

$$\begin{aligned} w_i - w_j &= w_i^{(g-1)/g} a_{ig}^{1/g} - w_j^{(g-1)/g} a_{jg}^{1/g} = w_i^{(g-1)/g} (a_{ij} a_{jg})^{1/g} - w_j^{(g-1)/g} a_{jg}^{1/g} = \\ &= a_{ig}^{1/g} \left[w_i^{(g-1)/g} a_{ij}^{1/g} - w_j^{(g-1)/g} \right] = a_{ig}^{1/g} \left[w_i^{(g-1)/g} a_{ij}^{1/g} - w_j^{(g-1)/g} \right] \geq 0. \end{aligned}$$

Справедливость последнего неравенства обеспечивается предположением $a'_{ij} = w'_i / w'_j$ и $w'_i > w'_j$.

В общем случае используются непосредственно формулы для вычисления ВКА по методу ММАИ (этапы II-IV). Осуществляя последовательно несложные, но громоздкие преобразования, можно показать, что и в общем случае $w_i > w_j$. Теорема 3.3 доказана.

Доказательство теоремы 3.4.

Функция относительного сходства имеет вид:

$$F(i) = g^{-1} \sum_{j=1}^g w_{ij}(i) \left(F'(i) = \left[\prod_{j=1}^g w_{ij}(i) \right]^{1/g} \right),$$

$$w_{ij}(i) = \sum_{s=1}^v c_s w_{ij}^s(i) = \sum_{s=1}^v c_s \frac{w_i^s}{w_i^s + w_j^s}, \quad w_i^s = \left(\prod_{l=1}^g a_{il}^s \right)^{1/g} \left[\sum_{k=1}^g \left(\prod_{l=1}^g a_{kl}^s \right)^{1/g} \right]^{-1},$$

величины $a_{kl}^s = w_k^s / w_l^s$ - элементы МПС по s -му критерию МАИ. Далее без ограничения общности доказательство проведем для $F(i) = g^{-1} \sum_{j=1}^g w_{ij}(i)$.

Обозначим через $i^* \in Y$ произвольную точку максимума функции относительного сходства на множестве выбираемых решений Y (номеров заданных альтернатив-признаков), $F(i^*) \geq F(i)$ для любого $i \in Y$. Перепишем выражение для $F(i)$ в виде:

$$F(i) = g^{-1} \sum_{j=1}^g w_{ij}(i) = g^{-1} \sum_{j=1}^g \sum_{s=1}^v c_s w_{ij}^s(i) = g^{-1} \sum_{s=1}^v c_s \sum_{j=1}^g w_{ij}^s(i)$$

Обозначим $f_s(i) = \sum_{j=1}^g w_{ij}^s(i)$ и установим парето-оптимальность точки $i^* \in Y$ [126, 133]. Доказательство этого факта проведем от противного. Пусть существует такая точка $i' \in Y$, что имеет место векторное неравенство: $f_s(i') \geq f_s(i^*)$. В силу справедливости аксиом выбора 1-3 [133] и определения парето-оптимальности последнее неравенство можно переписать покомпонентно в виде: $f_l(i') \geq f_l(i^*)$, для всех $l = \overline{1, v}$, при этом одно из неравенств должно быть строгим. Умножая обе части неравенства на постоянные положительные величины c_1, c_2, \dots, c_v $\left(\sum_i c_i = 1 \right)$ (последовательно, первое – на $c_1 > 0$, второе – на $c_2 > 0$ и т.д.) и складывая, получим неравенство, несовместное с предположением: $F(i') \geq F(i^*)$. Отсюда следует, что точка $i^* \in Y$ принадлежит парето-оптимальному множеству ($i^* \in P_f(Y)$). Теорема 3.4 доказана.

ПРИЛОЖЕНИЕ к главе 4

Доказательство теоремы 4.1.

Пусть заданы матрицы Q, R и T. Не умаляя общности, проведем доказательство для теста $\tau = (z_1, z_2, \dots, z_g)$. Обозначим через σ_i – количество различных i -м признаком различий типа «образ-образ» («объект-образ», «объект-объект»); через $\bar{\sigma}$ – количество всевозможных различий объектов из разных образов по матрице описаний Q. Тогда нормализованный коэффициент \bar{w}_m^{-2} по формуле (4.9) имеет вид:

$$\bar{w}_m^{-2} = \frac{\sigma_m}{\bar{\sigma} \cdot \sum_{j=1}^g (\sigma_j / \bar{\sigma})} = \frac{\sigma_m}{\sum_{j=1}^g \sigma_j}.$$

Поставим данному признаку z_m в соответствие совокупность различных m -м признаком пар объектов из разных образов и будем говорить, что признак z_m порождает мультимножество P_m .

Сформируем матрицу парных сравнений признаков, входящих в тест $\tau = (z_1, z_2, \dots, z_g)$ с мерой относительной важности признака i над признаком j ,

равной $\frac{|P_i|}{|P_j|}$ (рис. П4.1 а)), заметив, что σ_m интерпретируется в терминах мультимножеств как мощность P_m , т. е. $\sigma_m = |P_m|$. Далее получим оценку собственного вектора [141] (рис. П4.1)) и нормализованное его значение. Согласно методу МАИ и утверждению о приближении собственного вектора среднегеометрическим [141] значением по строкам МПС, будем иметь представление:

$$\begin{array}{ccc}
 \begin{array}{c} z_1 \\ z_2 \\ \dots \\ z_g \end{array} & \begin{array}{c} z_1 \quad z_2 \quad \dots \quad z_g \\ \left(\begin{array}{cccc} 1 & \frac{a_1}{a_2} & \dots & \frac{a_1}{a_g} \\ \frac{a_2}{a_1} & 1 & \dots & \frac{a_2}{a_g} \\ \dots & \dots & \dots & \dots \\ \frac{a_g}{a_1} & \frac{a_g}{a_2} & \dots & 1 \end{array} \right) \end{array} & \begin{array}{c} \left(\begin{array}{ccc} 1 \cdot \frac{a_1}{a_2} \cdot \frac{a_1}{a_3} \dots \frac{a_1}{a_g} \\ \frac{a_2}{a_1} \cdot 1 \cdot \frac{a_2}{a_3} \dots \frac{a_2}{a_g} \\ \dots \\ \frac{a_g}{a_1} \cdot \frac{a_g}{a_2} \dots \frac{a_g}{a_{g-1}} \cdot 1 \end{array} \right)^{1/g} & \left(\frac{\prod_{j=2}^g \frac{a_1}{a_j}}{\sum_{j=1}^g \left(\frac{a_j^{g-1}}{\prod_{k=1, k \neq j}^m a_k} \right)} \right)^{1/g} \\ \left(\frac{\prod_{j=1, j \neq 2}^g \frac{a_2}{a_j}}{\sum_{j=1}^g \left(\frac{a_j^{g-1}}{\prod_{k=1, k \neq j}^g a_k} \right)} \right)^{1/g} & \dots \\ \left(\frac{\prod_{j=1, j \neq g}^g \frac{a_g}{a_j}}{\sum_{j=1}^g \left(\frac{a_j^{g-1}}{\prod_{k=1, k \neq j}^g a_k} \right)} \right)^{1/g} & \dots \end{array} \\
 & \text{a)} & \text{б)} \qquad \text{c)}
 \end{array}$$

Рис. П4.1. Матрица парных сравнений признаков а); оценка собственного вектора б); нормализованное значение собственного вектора с)

Таким образом, значение весового коэффициента $w_{1,m}^6$ равно следующему выражению ($a_m = \sigma_m$):

$$w_{1,m}^6 = \frac{\left(\frac{\prod_{j=1, j \neq m}^g a_m}{\prod_{j=1, j \neq m}^g a_j} \right)^{1/g}}{\sum_{j=1}^g \left(\frac{a_j^{g-1}}{\prod_{k=1, k \neq j}^g a_k} \right)^{1/g}} = \frac{\left(a_m^{g-1} \right)^{1/g} / \left(\frac{\prod_{j=1, j \neq m}^g a_j}{\prod_{j=1, j \neq m}^g a_j} \right)^{1/g}}{\sum_{j=1}^g \left(a_j^g \right)^{1/g} / \left(\frac{\prod_{k=1}^g a_k}{\prod_{k=1}^g a_k} \right)^{1/g}} = \frac{a_m}{\sum_{j=1}^g a_j} = w_m^2.$$

Теорема 4.1 доказана.

Доказательство теоремы 4.2.

Доказательство опирается на лемму из работы [126].

1) Докажем выполнение свойств (3.1) для МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$, элементы которой построены по формулам (4.11) и (4.12). Справедливость первого и второго свойств (3.1) МПС $\mathbf{A} = \|a_{ij}\|_{M \times M}$, следует непосредственно из представле-

ния (4.11). Справедливость свойства совместности 3) (3.1) следует из (4.11) и представления (4.12), так как для всех номеров $i, j, k = \overline{1, M}$, выполняется:

$$a_{ik} \cdot a_{kj} = \frac{a_{1k} a_{1j}}{a_{1i} a_{1k}} = a_{ij}.$$

Справедливость свойства 4) свойств МПС (3.1) устанавливается использованием специального вида МПС A , характеристическое уравнение которой можно записать в виде [36, 37]:

$$\det(A - \lambda \cdot E) = \det \begin{bmatrix} (1 - \lambda) & 1 & \dots & 1 \\ 1 & (1 - \lambda) & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & (1 - \lambda) \end{bmatrix} = (-1)^M \lambda^{M-1} (\lambda - M) = 0,$$

из которого следует, что матрица A имеет только два собственных значения, при этом, значение M является максимальным.

2) Доказательство формулы (4.13) для компонент весового ненормализованного вектора $W_1^7 = (w_{11}^7, w_{12}^7, \dots, w_{1M}^7)^T$ проверяется непосредственно подстановкой $w_{1i}^7 = \frac{a_{iM}}{a_{1i}}$, $i = \overline{1, M}$ в равенство $A \cdot W_1^7 = M \cdot W_1^7$, учитывая справед-

ливость свойства совместности МПС A :

$$\sum_{i=1}^M a_{ki} \cdot w_{1i}^7 = \sum_{i=1}^M a_{ki} \cdot \frac{a_{iM}}{a_{1i}} = \sum_{i=1}^M a_{ki} \cdot a_{iM} = \sum_{i=1}^M a_{kM} = M \cdot a_{kM} = M \cdot w_{1k}^7, \quad k = \overline{1, M}.$$

Таким образом, для МПС $A = \|a_{ij}\|_{M \times M}$ с элементами, построенными по формулам (4.12) и (4.13), выполняется свойство совместности, ее единственность следует из (4.12), справедливость формулы (4.13) доказана. Теорема 4.2 доказана.

Доказательство теоремы 4.3.

Справедливость первых свойств 1) и 2) (3.1) следует непосредственно из (4.14) и (4.15). Справедливость свойства совместности 3) формулы (3.1) следует из (4.11) и представления (4.12), так как для всех номеров $i = \overline{1, M-2}$, $i < j-1$, $k = \overline{1, M}$ выполняется:

$$a_{ij} = a_{i, j-1} \cdot a_{j-1, j} = a_{i, j-2} \cdot a_{j-2, j} = \dots = a_{ik} \cdot a_{kj}.$$

Равенство $A \cdot W_2^7 = M \cdot W_2^7$ проверяется непосредственно подстановкой $w_{2i}^7 = a_{i, i+1} \cdot a_{i+1, i+2} \cdot \dots \cdot a_{M-1, M}$, $i = \overline{1, M-1}$, $w_M = 1$, с учетом справедливости свойства совместности $a_{ij} = a_{ik} \cdot a_{kj}$:

$$\sum_{i=1}^M a_{ki} \cdot w_{2i}^7 = \sum_{i=1}^M a_{ki} \cdot \prod_{j=i}^{M-1} a_{j,j+1} = \sum_{i=1}^M \prod_{j=k}^{M-1} a_{j,j+1} = M \cdot \prod_{j=k}^{M-1} a_{j,j+1} = M \cdot w_{2k}^7, \quad k = \overline{1, M}$$

Теорема 4.3 доказана.

Иллюстративный пример применения формализма мультимножеств

Иллюстрируем применение формализма мультимножеств при получении ВКП по методу 6, применяя при этом традиционный МАИ для простоты изложения (пример модификации ММАИ изложен в гл. 3). Методы 6 и 7 (гл. 4) на основе ММАИ и упрощенного МАИ реализованы в ИМСЛОГ. Изложение этапов определения ВКП построим на примере рис. 4.1.

Мультимножества, порожденные соответствующими признаками, имеют вид:

$$P_5 = \{4 \cdot (1-2), 1 \cdot (1-3), 2 \cdot (1-4), 8 \cdot (2-5), 4 \cdot (2-6), 2 \cdot (3-5), 1 \cdot (3-6), 4 \cdot (4-5), 2 \cdot (4-6)\};$$

$$P_9 = \{4 \cdot (1-2), 1 \cdot (1-6), 4 \cdot (2-3), 8 \cdot (2-4), 8 \cdot (2-5), 1 \cdot (3-6), 2 \cdot (4-6), 2 \cdot (5-6)\};$$

$$P_{11} = \{2 \cdot (1-4), 2 \cdot (1-5), 8 \cdot (2-4), 8 \cdot (2-5), 2 \cdot (3-4), 2 \cdot (3-5), 2 \cdot (4-6), 2 \cdot (5-6)\};$$

Запись $8 \cdot (2-5)$ означает, что элемент вида $(2-5)$ в мультимножество P_5 входит 8 раз. Мощности мультимножеств равны: $|P_5|=28$, $|P_9|=30$, $|P_{11}|=28$.

На этапе 1 составляем МПС признаков на основе первой меры относительной важности формул (4.11) (табл. П8.1).

Таблица П8.1. Весовые коэффициенты признаков на этапе 1 метода 6

	z_5	z_9	z_{11}	W_1^{dB}
z_5	1	28/30	1	0.326
z_9	30/28	1	30/28	0.349
z_{11}	1	28/30	1	0.326

На втором этапе вычисляем разности соответствующих мультимножеств и значения их мощностей:

$$P_5 - P_9 = \{1 \cdot (1-3), 2 \cdot (1-4), 4 \cdot (2-6), 2 \cdot (3-5), 4 \cdot (4-5)\};$$

$$P_9 - P_5 = \{1 \cdot (1-6), 4 \cdot (2-3), 8 \cdot (2-4), 2 \cdot (5-6)\};$$

$$P_5 - P_{11} = \{4 \cdot (1-2), 1 \cdot (1-3), 4 \cdot (2-6), 1 \cdot (3-6), 4 \cdot (4-5)\};$$

$$P_{11} - P_5 = \{2 \cdot (1-5), 8 \cdot (2-4), 2 \cdot (3-4), 2 \cdot (5-6)\};$$

$$P_9 - P_{11} = \{4 \cdot (1-2), 1 \cdot (1-6), 4 \cdot (2-3), 1 \cdot (3-6)\};$$

$$P_{11} - P_9 = \{2 \cdot (1-4), 2 \cdot (1-5), 2 \cdot (3-4), 2 \cdot (3-5)\}.$$

Таким образом, искомые величины равны:

$$|P_5 - P_9|=13, |P_9 - P_5|=15, |P_5 - P_{11}|=14, |P_{11} - P_5|=14, |P_9 - P_{11}|=10, |P_{11} - P_9|=8.$$

Составляем МПС на основе второй меры относительной важности (4.11) (табл. П8.2):

Таблица П8.2. Весовые коэффициенты признаков на этапе 2 метода б

	z_5	z_9	z_{11}	W_2^6
z_5	1	13/15	14/14	0.316
z_9	15/13	1	5/4	0.375
z_{11}	14/14	4/5	1	0.308

На третьем этапе вычисляем размерности соответствующих мультимножеств и значения их мощностей: $|P_5-P_9|=5$, $|P_9-P_5|=5$, $|P_5-P_{11}|=5$, $|P_{11}-P_5|=4$, $|P_9-P_{11}|=4$, $|P_{11}-P_9|=4$. Составляем МПС признаков на основе третьей меры относительной важности (4.11) (табл. П8.3):

Таблица П8.3. Весовые коэффициенты признаков на этапе 3 метода б

	z_5	z_9	z_{11}	W_3^6
z_5	1	5/4	5/4	0.385
z_9	4/5	1	1	0.308
z_{11}	4/5	1	1	0.308

Найдем обобщенные значения ВКП, входящих в тест, как среднее геометрическое величин, найденных на этапах 1 – 3 (табл. П8.4).

Таблица П8.4. Весовые коэффициенты (глобальные) признаков по методу б

	W_1^6	W_2^6	W_3^6	W^6
z_5	0.326	0.316	0.385	0.341
z_9	0.349	0.375	0.308	0.343
z_{11}	0.326	0.308	0.308	0.314

Алгоритм⁴⁴ П4.1. Поиск всех кратчайших столбцовых покрытий двоичной матрицы

Вход: двоичная матрица U размерности nm , где n - количество строк, m - количество столбцов.

Выход: неповторяющиеся кратчайшие столбцовые покрытия.

1: В заданной двоичной матрице U ищутся строки, содержащие ровно одну единицу. Соответствующие этим единицам столбцы включаются в текущее по-

⁴⁴ Автор алгоритма П4.1 - Янковская А.Е.

крытие и отмечаются символом «*». Если все столбцы в матрице U окажутся отмеченными, то процесс прекращается, поскольку найдено единственное кратчайшее столбцовое покрытие.

2: В текущей подматрице (матрице U) для выделения очередной подматрицы выбирается, отмечается символом «*» и включается в текущее покрытие еще неотмеченный столбец, имеющий наибольшее (если таковых несколько, то самый левый из них) количество единичных значений. Если мощность текущего покрытия превысила мощность ранее найденных покрытий, то переход к *шагу 6*.

3: Выделяется (и становится текущей) очередная подматрица по отмеченному символом «*» столбцу с одновременным удалением столбцов, отмеченных символом «+». Если в текущей подматрице есть нулевые столбцы, т.е. не покрывающие ни одной строки, то они отмечаются символом «-«. Если количество строк в текущей подматрице больше 1 (концевые вершины еще не достигнуты), то переход к *шагу 1*.

4: Если в текущей подматрице (строке) число неотмеченных столбцов больше 1, то левый из них отмечается символом «*» и включается в текущее покрытие, а остальные отмечаются символом «+». Если мощность текущего покрытия превысила мощность ранее найденных покрытий, то переход к *шагу 6*.

5: Если мощность текущего покрытия меньше мощности ранее найденных покрытий, то все они отбрасываются. Текущее покрытие запоминается. Если есть столбцы, отмеченные символом «+», то формируются равномошные столбцовые покрытия путем поочередной замены столбца, отмеченного символом «*», на каждый из отмеченных символом «+».

6: Восстанавливается (вместе с разметкой столбцов и текущим покрытием) и становится текущей подматрица, соответствующая в дереве поиска вершине предыдущего уровня. Если в текущей подматрице есть столбцы, отмеченные символом «+», то создаются равномошные столбцовые покрытия путем поочередной замены столбца, отмеченного символом «*», на каждый из отмеченных символом «+». Столбцы, отмеченные символами «*» и «+», переотмечаются символом «-«. Если в текущей подматрице (матрице U) есть неотмеченные столбцы, то переход к *шагу 1*.

7: Если текущая подматрица не является матрицей U , то переход к *шагу 6*, в противном случае процесс завершается.

Комментарий. Поиск всех кратчайших столбцовых покрытий двоичной матрицы U сводится к построению по матрице U иерархической системы подматриц, являющейся деревом поиска, и нахождению всех неповторяющихся кратчайших путей в дереве поиска. Для ускоренного поиска всех неповторяющихся кратчайших путей (кратчайших столбцовых покрытий) применяется специальная разметка вершин, позволяющая сократить число переборных и заключающаяся в том, что на каждом уровне дерева поиска отмечаются вершины, соответствующие нулевым и равным столбцам и столбцам, по которым уже выде-

лялись подматрицы. Учет первых позволяет отсекаать ветви, не соответствующие столбцовым покрытиям; вторых - находить столбцовые покрытия без выделения подматриц; третьих - исключить построение одинаковых столбцовых покрытий.

Правило выделения очередной подматрицы по некоторому столбцу сводится к вычеркиванию данного столбца и всех строк, содержащих единичное значение в этом столбце.

Корню древовидной структуры (дереву поиска) сопоставляется матрица U . Вершинам 1-го уровня сопоставляются подматрицы, выделенные из матрицы U . Вершинам 2-го уровня сопоставляются подматрицы, выделенные из подматриц 1-го уровня, и т.д. Наконец, конечным вершинам (листьям) сопоставляются вырожденные подматрицы, состоящие из одного столбца.

Напомним, что последовательность, состоящая из k вершин, соединяющих корень с концевой вершиной, называется путем длины k , а путь, состоящий из минимального числа вершин - кратчайшим.

В процессе построения дерева поиска каждой вершине приписывается значение, равное номеру столбца, по которому была выделена сопоставленная данной вершине подматрица. Таким образом, множество значений вершин, входящих в кратчайший путь, задает множество столбцов, входящих в соответствующее кратчайшее столбцовое покрытие двоичной матрицы.

ПРИЛОЖЕНИЕ к главе 5

Доказательство теоремы 5.2.

В соответствии с теоремой 5.1 для функции $H(P_\Omega, P_G)$ с вектором вероятностей $P_\Omega = (p^1, \dots, p^I)$ справедливы соотношения (5.4) (строгие неравенства соответствуют только нулевым значениям p^i).

Функция $-H(P_\Omega, P_G)$ является выпуклой кверху по построению (при $\hat{p}_v = \sum_{i=1}^I p^i p_v^i$ совпадает с функцией энтропии по Шеннону). Условия теоремы

5.1 выполнены в условиях теоремы 5.2, и соотношения (5.4) примут вид:

$$-\partial H(P_\Omega, P_G) / \partial p_i = \log_2 e \sum_{v \in U} \hat{p}_v p_v^i \left(\sum_{i=1}^I p^i p_v^i \right)^{-1} \leq \lambda, \quad (\text{П5.1})$$

Домножим каждую часть неравенства (П5.1) на p_i и просуммировать по i ($i = \overline{1, I}$). В результате получим выражение для левой части (П5.1):

$$\log_2 e \sum_{i=1}^I p^i \sum_{v \in U} \hat{p}_v \frac{P_v^i}{\sum_{i=1}^I p_i P_v^i} = \log_2 \sum_{v \in U} \hat{p}_v \frac{\sum_{i=1}^I p^i P_v^i}{\sum_{i=1}^I p^i P_v^i} = \log_2 \sum_{v \in U} \hat{p}_v = \log_2 e, \quad (\text{П5.2})$$

и для правой части (П5.1): $\lambda \sum_{i=1}^I p^i = \lambda$, что приводит к справедливости соотношения:

$$\log_2 e \sum_{v \in U} \hat{p}_v P_v^i \left(\sum_{i=1}^I p^i P_v^i \right)^{-1} = \log_2 e, \text{ или утверждению теоремы 5.2.}$$

Для доказательства единственности набора $P_\Omega = (p_1, \dots, p_I)$ предположим противное, а именно, существование двух наборов P_Ω, P'_Ω , при которых достигается $\min_{P_\Omega} H(P_\Omega, P_G) = \min_{P'_\Omega} H(P_\Omega, P_G) = H(P_G)$. Возьмем произвольное число $0 < \varepsilon < 1$, и для выпуклой по P_Ω (P'_Ω) функции $H(P_\Omega, P_G)$, применим неравенство Йенсена:

$$\begin{aligned} H(\varepsilon P_\Omega + (1-\varepsilon)P'_\Omega, P_G) &\leq \varepsilon H(P_\Omega, P_G) + (1-\varepsilon)H(P'_\Omega, P_G) = \\ &= \varepsilon H(P_G) + (1-\varepsilon)H(P_G) = H(P_G). \end{aligned} \quad (\text{П5.3})$$

Предположим, что наборы P_Ω, P'_Ω различаются по l -й компоненте. Из формулы (5.3) следует, что

$$H(P_\Omega, P_G) = -\log_2 p_y(P_\Omega)/n = -\sum_{v \in U} \hat{p}_v \log_2 \sum_{i=1}^I p^i P_v^i = -\sum_{v \in U} \log_2 \left(\sum_{i=1}^I p^i P_v^i \right)^{\hat{p}_v}$$

Рассмотрим выражение:

$$H(\varepsilon P_\Omega + (1-\varepsilon)P'_\Omega, P_G) = -\sum_{v \in U} \log_2 \left(\sum_{i=1}^I (\varepsilon p^i + (1-\varepsilon)p'^i) P_v^i \right)^{\hat{p}_v}.$$

Для отдельного слагаемого этой суммы справедливы неравенства:

$$\begin{aligned} -\log_2 \left(\sum_{i=1, i \neq l}^I p^i P_v^i + (\varepsilon p^l + (1-\varepsilon)p'^l) P_v^l \right)^{\hat{p}_v} &\leq -\log_2 \left((\varepsilon p^l + (1-\varepsilon)p'^l) P_v^l \right)^{\hat{p}_v} \leq \\ &\leq -\hat{p}_v \varepsilon \log_2 (p^l P_v^l) - \hat{p}_v (1-\varepsilon) \log_2 (p'^l P_v^l). \end{aligned}$$

С учетом полученных соотношений неравенство (П5.3) будет строгим, что противоречит предположению о минимальности $H(P_G)$. Теорема 5.2 доказана.

Доказательство леммы 5.1.

Содержательный смысл преобразования (5.6) следующий. Из формул

Байеса и полной вероятности следует, что

$$P(\Omega_i / \nu) = \frac{P(\Omega_i)P(\nu / \Omega_i)}{\sum_{j=1}^I P(\Omega_j)P(\nu / \Omega_j)}, \quad P(\Omega_i) = \sum_{\nu \in U} P(\nu)P(\Omega_i / \nu),$$

или в наших обозначениях:

$$P(i / \nu) = p^i p_\nu^i \left(\sum_{j=1}^I p^j p_\nu^j \right)^{-1}, \quad p^i = \sum_{\nu \in U} p_\nu p^i = \sum_{\nu \in U} p_\nu p^i p_\nu^i \left(\sum_{j=1}^I p^j p_\nu^j \right)^{-1}. \quad (\text{П5.4})$$

При условии $p^i > 0$ поделив последнее равенство в (П5.4) на p^i , получим уравнение правдоподобия (5.5), а в условиях теоремы 5.2 набор $P_\Omega = (p^1, \dots, p^I)$, обеспечивающий выполнение (П5.4), единственный и является ММП-оценкой искомого распределения состояний ДО.

Для доказательства 1-го утверждения леммы 5.1 достаточно показать выполнение условия нормировки:

$$\sum_{i=1}^I p^i = \sum_{i=1}^I \sum_{\nu \in U} \frac{\hat{p}_\nu p^i p_\nu^i}{\sum_{i=1}^I p^i p_\nu^i} = \sum_{i=1}^I p^i \sum_{\nu \in U} \hat{p}_\nu \frac{p_\nu^i}{\sum_{i=1}^I p^i p_\nu^i} = \sum_{\nu \in U} \hat{p}_\nu \frac{\sum_{i=1}^I p^i p_\nu^i}{\sum_{i=1}^I p^i p_\nu^i} = \sum_{\nu \in U} \hat{p}_\nu = 1. \quad (\text{П5.5})$$

Для доказательства 2-го утверждения леммы 5.1 покажем выполнение и неравенства:

$$\begin{aligned} p^i &= p^i \sum_{\nu \in U} \hat{p}_\nu p_\nu^i \left(\sum_{i=1}^I p^i p_\nu^i \right)^{-1} = p^i \left(\sum_{\nu \in \Omega_i} \hat{p}_\nu p_\nu^i \left(\sum_{i=1}^I p^i p_\nu^i \right)^{-1} + \sum_{\nu \notin \Omega_i} \hat{p}_\nu p_\nu^i \left(\sum_{i=1}^I p^i p_\nu^i \right)^{-1} \right) = \\ &= p^i \left(1 + \sum_{\nu \notin \Omega_i} \hat{p}_\nu p_\nu^i \left(\sum_{i=1}^I p^i p_\nu^i \right)^{-1} \right) \geq p^i. \end{aligned}$$

Лемма 5.1 доказана.

Доказательство леммы 5.2.

Рассмотрим разность $h(P_i^i, P_{i-1}^i)$:

$$\begin{aligned} h(P_i^i, P_{i-1}^i) &= H(P_i^i, P_G) - H(P_{i-1}^i, P_G) = -\sum_{\nu \in U} \hat{p}_\nu \log_2 \sum_{i=1}^I p^i p_\nu^i + \sum_{\nu \in U} \hat{p}_\nu \log_2 \sum_{j=1}^I p^j p_\nu^j = \\ &= -\sum_{\nu \in U} \hat{p}_\nu \left(-\log_2 \sum_{j=1}^I p^j p_\nu^j + \log_2 \sum_{i=1}^I p^i p_\nu^i \right) = -\sum_{\nu \in U} \hat{p}_\nu \left(\log_2 \sum_{i=1}^I p^i p_\nu^i \left(\sum_{j=1}^I p^j p_\nu^j \right)^{-1} \right). \end{aligned}$$

Подставим в последнее выражение p^i из (5.6):

$$\begin{aligned}
h(P'_\Omega, P_\Omega) &= -\sum_{v \in U} \hat{p}_v \left(\log_2 \sum_{i=1}^I \sum_{s \in U} \frac{\hat{p}_s p^i p_s^i}{\sum_{l=1}^I p^l p_s^l} p_v^i \left(\sum_{j=1}^I p^j p_v^j \right)^{-1} \right) = \\
&= -\sum_{v \in U} \hat{p}_v \left(\log_2 \sum_{i=1}^I \sum_{s \in U} \frac{\hat{p}_s p_s^i}{\sum_{l=1}^I p^l p_s^l} \frac{p^i p_v^i}{\sum_{j=1}^I p^j p_v^j} \right) \leq -\sum_{v \in U} \hat{p}_v \left(\sum_{i=1}^I \frac{p^i p_v^i}{\sum_{j=1}^I p^j p_v^j} \log_2 \sum_{s \in U} \frac{\hat{p}_s p_s^i}{\sum_{l=1}^I p^l p_s^l} \right) = \\
&= -\sum_{i=1}^I \left(\sum_{v \in U} \hat{p}_v \frac{p^i p_v^i}{\sum_{j=1}^I p^j p_v^j} \right) \log_2 \left(\sum_{s \in U} \frac{\hat{p}_s p_s^i}{\sum_{l=1}^I p^l p_s^l} \right).
\end{aligned}$$

Справедливость неравенства следует из выпуклости функции $-\log_2 x$ и неравенства Йенсена, условие для применения которого в данном случае сводится к требованию $\sum_{i=1}^I p^i p_v^i \left(\sum_{j=1}^I p^j p_v^j \right)^{-1} = 1$, выполненного в силу (П5.5).

Преобразуем последнее выражение, переставив пределы суммирования и, воспользовавшись выражением (5.6):

$$\begin{aligned}
h(P'_\Omega, P_\Omega) &\leq -\sum_{i=1}^I \left(\sum_{v \in U} \hat{p}_v \frac{p^i p_v^i}{\sum_{j=1}^I p^j p_v^j} \right) \log_2 \left(\sum_{s \in U} \frac{\hat{p}_s p_s^i}{\sum_{l=1}^I p^l p_s^l} \right) = \\
&= -\sum_{i=1}^I p^i \log_2 \left(\frac{p^i}{p^i} \right) = -\sum_{i=1}^I p^i \log_2 p^i + \sum_{i=1}^I p^i \log_2 p^i.
\end{aligned} \tag{П5.6}$$

На основании леммы 5.1 следует справедливость неравенства:

$$-\sum_{i=1}^I p^i \log_2 p^i + \sum_{i=1}^I p^i \log_2 p^i \leq 0 \text{ и } h(P_l^i, P_{l-1}^i) < 0,$$

равенство возможно только при совпадении $\{p^i, i = \overline{1, I}\}$ и $\{p^i, i = \overline{1, I}\}$.

Лемма 5.2 доказана.

Доказательство теоремы 5.3.

Пусть \hat{P} - предельная точка последовательности $\{\hat{P}_l, l \geq 1\}$, к которой сходится и подпоследовательность $\{\hat{P}_k, k \geq 1\}$, при этом точка \hat{P} принадлежит

замкнутому множеству Θ . Тогда подпоследовательность $\{\hat{P}_{l_k+1}, k \geq 1\}$ сходится к точке $F(\hat{P})$ по построению (5.9), а величина $H(\hat{P}_{l_k+1}, P_G)$ сходится к точке $H(F(\hat{P}), P_G)$ как не возрастающая по l .

С другой стороны, вследствие не возрастания величины $H(\hat{P}_l, P_G)$ по l и сходимости $H(\hat{P}_k, P_G)$ к величине $H(\hat{P}, P_G)$, к этой же величине сходится и последовательность $H(\hat{P}_l, P_G)$. Таким образом, $H(F(\hat{P}), P_G) = H(\hat{P}, P_G)$, или точка \hat{P} является неподвижной точкой преобразования (5.9), единственной согласно (5.10) и теоремы 5.2. Теорема 5.3. доказана.

ПРИЛОЖЕНИЕ к главе 6

Доказательство леммы 6.3

Из (6.2) следует, что

$$\varphi_j'(a_j) = M \zeta_i \exp(\lambda \zeta_i - \varphi_i(\lambda)) \quad (\text{П6.1})$$

$$\varphi_j''(a_j) = M(\zeta_i - \varphi_i'(\lambda))^2 \exp(\lambda \zeta_i - \varphi_i(\lambda)), \quad i = \overline{1, n}. \quad (\text{П6.2})$$

Так как $\varphi_j''(a_j) \geq 0$, то $\varphi_i'(\lambda)$ не убывает по λ . Учитывая, что $\varphi_i'(0) = 0$, функция $\varphi_i'(\lambda) \leq 0$ при $\lambda < 0$. Из леммы 2 следует, что $|\varphi_i'(\lambda)| \geq \varepsilon_c \cdot \sigma_i^2 \cdot L_i$, $\varepsilon_c = \ln c \cdot c^{-1} / 2$ для всех $|\lambda| \leq \ln c \cdot L_i^{-1} / 2 = \Lambda_i$, где $c > 1$, $|\zeta_i| \leq L_i$, $\sigma_i = D\zeta_i$. Поэтому в силу монотонного неубывания $\varphi_i'(\lambda)$ и вышеприведенного неравенства, последовательность $\{a_i\}$ будет удовлетворять уравнению (6.10), если его правая часть Δg_i удовлетворяет неравенству: $\kappa |q_0 m - \nu_i| \leq \varepsilon_c \cdot \sigma_i^2 \cdot L_i$. Лемма 6.3 доказана.

Доказательство леммы 6.4.

Рассмотрим $\tilde{M}_n \bar{\Phi}_n$:

$$\begin{aligned} \tilde{M}_n \bar{\Phi}_n &= M \bar{\Phi}_n G_n = \sum_{i=1}^n [M \zeta_i \exp(a_i \zeta_i - \varphi_i(a_i)) \times \\ &\times \exp(\sum_{j=1, j \neq i}^n (a_j \zeta_j - \varphi_j(a_j))) + (\Delta g_i - \beta_i) G_r] = \\ &= \sum_{i=1}^n [M \zeta_i \exp(a_i \zeta_i - \varphi_i(a_i)) \cdot M \exp(\sum_{j=1, j \neq i}^n (a_j \zeta_j - \varphi_j(a_j))) - \Delta g_i G_n] = \end{aligned}$$

$$= \sum_{i=1}^n \left[\varphi_i'(a_i) - \Delta g_i G_n \right] = 0.$$

Последнее равенство справедливо в силу леммы 6.3.

Аналогично проверяется мартингалное свойство:

$$\begin{aligned} \tilde{M}_n \left[\bar{\Phi}_n / \mathcal{F}_{n-1} \right] &= \tilde{M}_n \left[\bar{\zeta}_n / \mathcal{F}_{n-1} \right] + \tilde{M}_n \left[\bar{\Phi}_{n-1} / \mathcal{F}_{n-1} \right] = \bar{\Phi}_{n-1} + \tilde{M}_n \bar{\zeta}_n = \\ &= \bar{\Phi}_{n-1} + M \bar{\zeta}_n G_n = \bar{\Phi}_{n-1}. \end{aligned}$$

Лемма 6.4 доказана.

Доказательство теоремы 6.1.

Положим в (6.3) $A = \{\tau > n\}$ и, учитывая, что $G_n > 0$ (P-п.н. и \tilde{P}_n -п.н.), перепишем (6.3) в виде:

$$P\{\tau > n\} = \tilde{M}_n \chi_A \cdot G_n^{-1}, \quad A = \{\tau > n\}, \quad (\text{П6.3})$$

где χ_A – индикатор множества A .

Применим к (П6.3) неравенство Гельдера:

$$P\{\tau > n\} \leq \tilde{P}_n^{1/p} \{\tau > n\} \cdot \tilde{M}_n^{1/q} G_n^{-1}, \quad p, q > 1. \quad (\text{П6.4})$$

Оценим $\tilde{M}_n^{1/q} G_n^{-1}$:

$$\begin{aligned} \tilde{M}_n^{1/q} G_n^{-q} &= M^{1/q} G_n^{1-q} = \exp \sum_{i=1}^n \left(\frac{1}{q} \varphi_i'((1-q)a_i) + \frac{q-1}{q} \varphi_i'(a_i) \right) \leq \\ &\leq \exp \sum_{i=1}^n \left(\frac{1}{q} (1-q)^2 c^2 \sigma_i^{-2} \left(\varphi_i'(\lambda) \right)^2 + \frac{q-1}{q} c^2 \sigma_i^{-2} \left(\varphi_i'(\lambda) \right)^2 \right) = \\ &= \exp \sum_{i=1}^n \left(\frac{q-1}{q} c^2 \sigma_i^{-2} \left(\varphi_i'(a_i) \right)^2 \right) = \exp \sum_{i=1}^n \left((q-1) c^2 \sum_{i=1}^n \sigma_i^{-2} \left(\varphi_i'(a_i) \right)^2 \right). \end{aligned} \quad (\text{П6.5})$$

Справедливость неравенства следует из утверждения леммы 6.2, переписанном в виде: $|\lambda| c^{-1/2} \sigma_i^2 \leq \left| \varphi_i'(\lambda) \right| \leq c \cdot \sigma_i^2 |\lambda|$, откуда следует, что:

$$\begin{aligned} \left(\varphi_i'(\lambda) \right)^2 \sigma_i^{-2} c^{-1} &\leq |a_i| \leq c^{1/2} \sigma_i^{-2} \left(\varphi_i'(\lambda) \right)^2, \\ \left(\varphi_i'(\lambda) \right)^2 \sigma_i^{-2} c^{-5/2} &\leq \left| \varphi_i'(\lambda) \right| \leq c^2 \sigma_i^{-2} \left(\varphi_i'(\lambda) \right)^2, \end{aligned} \quad (\text{П6.6})$$

$$\left| \varphi_i'((1-q)a_i) \right| \leq c^2 (1-q)^2 \sigma_i^{-2} \left(\varphi_i'(\lambda) \right)^2, \quad \text{для всех } |(1-q)a_i| < \Lambda_i.$$

Оценим сверху вероятность $\tilde{P}_n\{\tau > n\}$, отметив, что $\tilde{P}_n\{\tau > n\} = \tilde{P}_n\{\tau_0 > n\} / 2$, где τ_0 – момент выхода из единичной полосы мартингала (μ_n, \mathcal{F}_n) , $\mu_n = \bar{S}_n / b$ т.е. $\tau_0 = \inf\{n \geq 1 : |\bar{S}_n / b| > 1\}$ (действительно,

событие $A = \{|\mu_n| > 1\} = A_1 + A_2$, где несовместные симметричные события A_1, A_2 определяются как $A_1 = \{\mu_n > 1\}$ и $A_2 = \{\mu_n < -1\}$, для чего воспользуемся леммой Блэкуэлла и Фридмана. Из теоремы Дуба следует, что возрастающий процесс для мартингала (μ_n, \mathcal{F}_n) есть величина $\langle \mu_n \rangle$, определяемая по формуле: $\langle \mu_n \rangle = \sum_{j=1}^n \tilde{M}_n \left[(\Delta \mu_j)^2 / \mathcal{F}_{j-1} \right] \sqrt{b^2 - 4ac}$. С учетом утверждения

лемм 6.1 и 6.3 справедливы соотношения (b – параметр из условий теоремы 6.1):

$$\begin{aligned} \langle \mu_n \rangle &= \frac{1}{b^2} \sum_{j=1}^n \tilde{M}_n \left[(\bar{\zeta}_j)^2 / \mathcal{F}_{j-1} \right] = \frac{1}{b^2} \sum_{j=1}^n \tilde{M}_n \left[(\zeta_j - \Delta g_j + b/n)^2 / \mathcal{F}_{j-1} \right] = \\ &= \frac{1}{b^2} \sum_{j=1}^n \tilde{M}_n \left[(\zeta_j - \varphi'_j(a_j))^2 / \mathcal{F}_{j-1} \right] = \frac{1}{b^2} \sum_{j=1}^n \varphi''_j(a_j). \end{aligned}$$

Применяя лемму Блэкуэлла и Фридмана с учетом оценки леммы 6.2 для величины $\varphi''_j(a_j)$, получаем:

$$\begin{aligned} \tilde{P}_n\{\tau > n\} &= \tilde{P}_n\{\tau_0 > n\} / 2 = \tilde{P}_n\{\langle \mu \rangle_{\tau_0} > \langle \mu \rangle_n\} / 2 \leq \tilde{P}_n\left\{ \langle \mu \rangle_{\tau_0} > \frac{1}{b^2} \sum_{j=1}^n \sigma_j^2 c^{-1/2} \right\} / 2 \leq \\ &\leq \frac{(5\alpha^{-1})^{1/2}}{2} \exp\left\{ -\frac{c^{-1/2}(\pi^2/8)(1-3\alpha^{1/2})}{b^2} \sum_{j=1}^n \sigma_j^2 \right\}. \end{aligned} \quad (П6.7)$$

Собирая неравенства (П6.4), (П6.5), (П6.7), получаем первое утверждение теоремы 6.1 (верхнюю оценку для вероятности $P\{\tau > n\}$).

$$\begin{aligned} P\{\tau > n\} &\leq \tilde{P}_n^{1/p}\{\tau > n\} \cdot M_n^{1/q} G_n^{-1} \leq \left(\frac{(5\alpha^{-1})^{1/2}}{2} \right)^{1/p} \exp\left\{ -\frac{c^{-1/2}(\pi^2/8)(1-3\alpha^{1/2})}{pb^2} \sum_{j=1}^n \sigma_j^2 \right\} \times \\ &\times \exp\left\{ \sum_{i=1}^n \left((q-1)c^2 \sum_{i=1}^n \sigma_i^{-2} (\varphi'_i(a_i))^2 \right) \right\} = \left(\frac{(5\alpha^{-1})^{1/2}}{2} \right)^{1/p} \times \\ &\times \exp\left\{ -\frac{c^{-1/2}(\pi^2/8)(1-3\alpha^{1/2})}{pb^2} \sum_{j=1}^n \sigma_j^2 + (q-1)c^2 \sum_{i=1}^n \sigma_i^{-2} (\varphi'_i(a_i))^2 \right\} \end{aligned}$$

Для получения нижней оценки вероятности $P\{\tau > n\}$ применим к правой части формулы (6.3) неравенство Гельдера:

$$\tilde{P}_n\{\tau > n\} \leq P^{1/p}\{\tau > n\} \cdot M^{1/q} G_n^q, \quad (П6.8)$$

и оценим величины в правой части неравенства сверху, применяя неравенства (П6.6):

$$M^{1/q} G_n^q = \exp \sum_{i=1}^n \left(\frac{1}{q} \varphi_i(qa_i) - \varphi_i(a_i) \right) \leq \exp \sum_{i=1}^n \left((cq - c^{-5/2} / 2) \sum_{i=1}^n \sigma_i^{-2} (\varphi_i'(a_i))^2 \right). \quad (\text{П6.9})$$

Применим лемму Блэкуэлла и Фридмана (с учетом оценки сверху для величины $\varphi_j''(a_j)$ (лемма 6.2)) для оценивания снизу величины $\tilde{P}_n\{\tau > n\}$:

$$\begin{aligned} \tilde{P}_n\{\tau > n\} &= \tilde{P}_n\{\tau_0 > n\} / 2 = \tilde{P}_n\{\langle \mu \rangle_{\tau_0} > \langle \mu \rangle_n\} / 2 \geq \\ &\geq \tilde{P}_n^{1/2} \left\{ \langle \mu \rangle_{\tau_0} > \frac{c}{b^2} \sum_{j=1}^n \sigma_j^2 \right\} / 2 \geq \exp \left\{ -\frac{\pi^2}{16} (1 + 15\alpha^{3/4}) \left(\alpha^2 + \frac{c}{b^2} \sum_{j=1}^n \sigma_j^2 \right) \right\}. \end{aligned} \quad (\text{П6.10})$$

Собирая неравенства (П6.8)–(П6.10), получаем второе утверждение теоремы 6.1 (нижнюю оценку для вероятности $P\{\tau > n\}$):

$$\begin{aligned} P\{\tau > n\} &\geq \tilde{P}_n\{\tau > n\} M^{-1/q} G_n^q \geq \exp \left\{ -\frac{p\pi^2}{16} (1 + 15\alpha^{3/4}) \left(\alpha^2 + \frac{c}{b^2} \sum_{j=1}^n \sigma_j^2 \right) \right\} \times \\ &\times \exp \left\{ -p(cq - c^{-5/2} / 2) \sum_{i=1}^n \sigma_i^{-2} (\varphi_i'(a_i))^2 \right\} = \\ &= \exp \left\{ -\frac{p\pi^2}{16} (1 + 15\alpha^{3/4}) \left[\alpha^2 - \frac{c}{b^2} \sum_{j=1}^n \sigma_j^2 \right] - p(cq - c^{-5/2} / 2) \sum_{i=1}^n \sigma_i^{-2} (\varphi_i'(a_i))^2 \right\}. \end{aligned}$$

Теорема 6.1 доказана.

Обозначения и сокращения

AR (AR-модель) – авторегрессия;
FRiS-функция (Function of Rival Similarity) – мера сходства объектов;
АД – асинхронный двигатель;
АлР – алгоритм(ы) распознавания;
АКАР – метод аналитического конструирования агрегированных регуляторов;
ВК, ВКА, ВКП – весовые коэффициенты, ВК альтернатив, ВК признаков;
ВР – временной ряд;
ДО – динамический объект;
ИМ – информационно-аналитическая интеллектуальная модель;
ИИС – интеллектуальное инструментальное средство
ИПП – информационное признаковое пространство;
ИТ – изображающая точка;
МАИ – метод анализа иерархий;
ММАИ – модифицированный метод анализа иерархий;
МНК – метод наименьших квадратов;
ОДЗ - область допустимых значений (центр ОДЗ);
ОДУ – обыкновенные дифференциальные уравнения;
с.в. – случайные величины;
СГК – система градаций квантования;
СММ – стохастическое мультимножество;
СКО – среднеквадратичное отклонение;
СОФ – сопровождающий оптимизирующий функционал;
СП – случайный процесс;
ЭК – элементарный классификатор;
ЭМО– электромеханический объект;
ЭМС – электромеханическая система;
ARADS - метод (алгоритм), от AutoRegression, Adaptive algorithm, Difference Scheme;
IReDSS – Intelligent software system of Recognition of Dynamic System State;
 $R (R_+)$ – поле вещественных чисел; $R_+ = \{x \in R \mid x \geq 0\}$;
 N – множество натуральных чисел;
 $Z (Z_+)$ – множество целых чисел ($Z_+ = \{x \in Z \mid x \geq 0\}$);
 A^T – «Т» означает транспонирование;
 R^n – линейное пространство над полем вещественных чисел
 $\|a\|$ – евклидова норма элемента $a \in R^n$;
 $\hat{m}(x)$ – обозначение оценки величины $m(x)$;
 $d(x, y), r(x, y)$ – функции расстояния.

Научное издание

БУКРЕЕВ Виктор Григорьевич
КОЛЕСНИКОВА Светлана Ивановна
ЯНКОВСКАЯ Анна Ефимовна

**ВЫЯВЛЕНИЕ ЗАКОНОМЕРНОСТЕЙ
ВО ВРЕМЕННЫХ РЯДАХ В ЗАДАЧАХ
РАСПОЗНАВАНИЯ СОСТОЯНИЙ
ДИНАМИЧЕСКИХ ОБЪЕКТОВ**

Монография

Издано в авторской редакции

Научный редактор *доктор технических наук,
профессор В.Г. Букреев*
Компьютерная верстка *С.И. Колесникова*
Дизайн обложки *О.Ю. Аршинова*


**Отпечатано в Издательстве ТПУ в полном соответствии
с качеством предоставленного оригинал-макета**

Подписано к печати 19.07.2010. Формат 60x84/16. Бумага «Снегурочка».
Печать XEROX. Усл. печ. л. 14,77. Уч.-изд. л. 13,36.
Заказ 1253-10. Тираж 500 экз.



Национальный исследовательский Томский политехнический университет
Система менеджмента качества
Томского политехнического университета сертифицирована
NATIONAL QUALITY ASSURANCE по стандарту ISO 9001:2008



ИЗДАТЕЛЬСТВО  ТПУ . 634050, г. Томск, пр. Ленина, 30
Тел./факс: 8(3822)56-35-35, www.tpu.ru