

Посвящается XX-летию конференции

Мягкие вычисления и измерения.

Теоретические основы и методы

Под редакцией проф. С.В. Прокопчиной

Том I



Москва

Издательский дом
«НАУЧНАЯ БИБЛИОТЕКА»

2017

УДК 519.6
ББК В 19л0 +3 965.07л0
М 43

РЕЦЕНЗЕНТЫ:

Кондрашкова Галина Анатольевна

*доктор технических наук, Санкт-петербургский национальный
исследовательский университет информационных технологий,
механики и оптики*

Котельников Валерий Григорьевич

*доктор технических наук, Финансовый университет
при Правительстве Российской Федерации*

АВТОРСКИЙ КОЛЛЕКТИВ:

С.В. Прокопчина, В.Б. Тарасов, В.Л. Лазарев,
В.Л. Горохов, Ю.В. Барышев, В.В. Витковский, М.С. Куприянов,
Л.Г. Логинская, Ю.А. Шичкина

Мягкие вычисления и измерения

М 43 **Теоретические основы и методы:** монография. Том I / под ред.
проф. С.В. Прокопчиной. – М.: ИД «НАУЧНАЯ БИБЛИОТЕКА»,
2017. – 420 с.

ISBN 987-5-9909931-7-4

Рассмотрены мягкие подходы к измерениям и управлению сложными системами, регулирующий байесовский подход, теоретический путь от нечетких множеств к мягким оценкам и синергетическому искусственному интеллекту, проанализированы методы построения гранулярных логических значений и структур, моделирования мнений и оценок интеллектуальных агентов: от четырехзначных систем модальностей к неклассическим мерам и нечетким распределениям. Рассмотрена теория энтропийных потенциалов, ее состояние, перспективы развития и практического приложения, методология когнитивной визуализации многомерных данных, нечеткие методы самоорганизации информационных систем, методы повышения эффективности генетических алгоритмов, общая концепция и подход к построению распределенных самоорганизующихся информационных систем, байесовские сети, родственные модели и их приложения.

Монография предназначена для научных работников, аспирантов, студентов и других специалистов работающих в сфере мягких вычислений и измерений, в области создания методов и средств искусственного интеллекта.

УДК 519.6

ББК В 19л0 +3 965.07л0

ISBN 987-5-9909931-7-4

© Коллектив авторов, 2017

© ИД «НАУЧНАЯ БИБЛИОТЕКА», 2017

ОГЛАВЛЕНИЕ

Предисловие	10
Глава 1. Мягкие подходы к измерениям и управлению сложными системами. Регуляризирующий Байесовский подход	12
1.1. Концептуальные основы решения задач в условиях неопределенности	12
1.1.1. Основные принципы решения задач измерения и управления сложными системами в условиях неопределенности	12
1.1.2. Обобщенная методика и схема решения задач в условиях неопределенности	19
1.1.3. Методологические основы регуляризирующего байесовского подхода для решения задач в условиях неопределенности	26
1.1.4. Модели и шкалы с динамическими ограничениями. Концепция, типы, структура	32
1.1.5. Основные принципы интеграции информации на основе регуляризирующего байесовского подхода (РБП) и его технологий	41
1.1.6. Метрологическое обоснование информационных технологий решения задач в условиях неопределенности	45
1.2. Организация измерительных процессов в условиях неопределенности. Регуляризирующий байесовский подход	47
1.2.1. Исторические аспекты развития теории измерений и интеллектуализации измерительных процессов	47
1.2.2. Основная концепция байесовской интеллектуализации измерений	55
1.2.3. Новые типы измерений в условиях неопределенности: нечеткие, мягкие, когнитивные измерения	56
1.2.4. Когнитивные измерения на основе байесовских интеллектуальных технологий	63

1.3. Интеллектуальные информационные технологии на основе регуляризирующего байесовского подхода	73
1.3.1. Интегрирующие технологии на основе байесовских интеллектуальных измерений	73
1.3.2. Принципы создания развивающихся информационных технологий на основе регуляризирующего байесовского подхода	76
1.3.3. Интеллектуальные технологии решения многокритериальных задач в условиях риска и неопределенности на основе регуляризирующего байесовского подхода	83
1.3.4. Методология и технологии мягких критериев и мягкого нормирования	92
1.3.5. Мягкие функции и их применение в управлении сложными системами	95
1.3.6. Методология и технологии байесовской математической статистики на основе РБП. Система математической обработки статистической информации «байесовская математическая статистика»	99
1.4. Мягкие технологии аудита и управления сложными системами в условиях неопределенности	106
1.5. Решение прикладных задач на основе регуляризирующего байесовского подхода	122
1.5.1. Опыт применения РБП и БИТ для решения измерительных и управленческих задач	122
1.5.2. Реализация принципов экономики знаний на основе БИИ	125
1.5.3. Решение социально-экономических задач на основе байесовских интеллектуальных технологий	131
1.5.4. Моделирование банковского сектора и показателей экономических систем ЕАЭС	133
1.5.5. Методологические аспекты создания интеллектуальных энергетических сетей (smart grid) .	137
Список литературы	142
Глава 2. От нечетких множеств к мягким оценкам и синергетическому искусственному интеллекту	145
Введение	145
2.1. НЕ-факторы, мягкие оценки и инженерия мнений	147
2.2. Лингвологический подход к формализации НЕ-факторов: принципы, этапы, модели	151

2.2.1. Моделирование нечеткости	152
2.2.2. Моделирование противоречивости	154
2.2.3. Моделирование неточности	155
2.3. От семейств НЕ-факторов к гибридным моделям и синергетическому ИИ	155
Список литературы	157
Глава 3. О методах построения гранулярных логических значений и структур	160
Введение	160
3.1. Гранулярные значения истинности и нестандартные семантики	161
3.1.1. Классические интерпретации значений истинности	161
3.1.2. Нестандартные логические семантики и обобщенные значения истинности	164
3.2. От логических миров к гранулярным логическим матрицам	166
3.2.1. Логические миры и пространства	166
3.2.2. Логические матрицы	168
Список литературы	171
Глава 4. Моделирование мнений и оценок интеллектуальных агентов: от четырехзначных систем модальностей к неклассическим мерам и нечетким распределениям	174
Введение	174
4.1. Мнения и модальности	176
4.1.1. Мнения	176
4.1.2. Описания и предписания в деятельности агента	178
4.1.3. Ценности, оценки и нормы	179
4.1.4. Единый подход к описанию различных систем модальностей на основе четырехзначных семантик	181
4.1.5. Решетки модальностей	183
4.2. Неклассические меры и описание модальностей	184
4.2.1. Вероятностная мера	184
4.2.2. Семейство мер Сугено	185
4.2.3. Основы теории свидетельств	186
4.2.4. Меры возможности и необходимости	187
4.2.5. Качественные оценки возможности и нечеткости ...	192
4.3. Меры на нечетких множествах	192
4.3.1. Меры энергии (показатели силы принадлежности) ..	193
4.3.2. Меры энтропии	194
4.3.3. Меры специфичности	194

Заключение	195
Список литературы	196
Глава 5. Теория энтропийных потенциалов. Состояние, перспективы развития и практического приложения	199
Введение	199
5.1. Отправные положения	200
5.2. Обзор подходов к описанию состояний неопределенности	202
5.3. Основы теории энтропийных потенциалов	204
5.4. Исследование систем на основе портретов в пространстве характеристик энтропийных потенциалов	213
5.5. Информационные портреты	217
5.6. Приложение методов ТЭП к решению практических задач	228
Заключение	234
Список литературы	235
Глава 6. Методология когнитивной визуализации многомерных данных	240
Введение	240
6.1. Когнитивные технологии как основа для нового поколения систем поддержки интеллектуальных решений	241
6.1.1. Современное состояние исследований по данной проблеме, основные направления исследований в мировой науке	241
6.1.2. Направление исследований	241
6.1.3. Список основных публикаций	244
6.2. Расширение и обновление поля рационального анализа данных на основе когнитивных интерфейсов	246
6.2.1. Механизмы визуализации многомерных данных	248
6.2.2. Пример когнитивной визуализации многомерных терабайтных данных астрофизического обзора COSMOS	253
6.3. Пример выявления аномальных ситуаций при мониторинге сложных систем на основе когнитивных технологий	254
6.3.1. Описание методики выявления аномалий	255
6.3.2. Выявление динамики	258
6.3.3. Выявление зависимостей по парам данных	259
6.3.4. Сравнение с другими методами	261
6.4. Технология и методика генерации псевдо 6-D образов ...	262

6.5. Выявление экстремальных состояний сложных природных и технических систем на основе когнитивного мониторинга многомерных временных рядов	265
6.5.1. Алгоритмы когнитивной визуализации многомерных временных рядов	266
6.5.2. Практическое применение когнитивных технологий для выявления аномальных состояний объектов мониторинга	267
6.6. Непараметрические статистики в многомерных временных рядах для когнитивного обнаружения аномалий	269
6.6.1. Этапы когнитивного анализа	269
6.6.2. Устойчивые (робастные) алгоритмы выявления аномалий	270
6.6.3. Практическое применение	271
6.7. Практика формирования инженерных визуальных метафор на основе когнитивных образов многомерных данных техносферного мониторинга	274
6.8. Когнитивные технологии визуализации многомерных данных для интеллектуальной поддержки принятия решений	276
6.8.1. Принципы работы	277
6.8.2. Новые алгоритмы когнитивной визуализации	280
6.9. Семиотика визуальных инженерных метафор когнитивных образов многомерных данных – новая парадигма компьютерной эргономики	281
Заключение	284
Список литературы	285
Глава 7. Нечеткие методы самоорганизации информационных систем	291
7.1. Нечеткие числа и логические значения	292
7.2. Семантические шкалы для измерения значений качественных понятий	296
7.3. Процедура порождения базиса нечетких термов, связанного с лингвистической парой	296
7.4. Семантические шкалы для измерения значений качественных понятий	297
7.5. Разложение нечеткого значения, связанного с лингвистической парой по базовым термам этой пары	299
7.6. Аппроксимация логических значений	300
7.7. Измерение качественных величин	301

7.8. Градуирование лингвистической шкалы	302
7.9. Арифметические операции над нечеткими числами	304
7.10. Логические переменные и операции	307
7.11. Аппроксимация логических значений	308
7.12. Логическая шкала как средство связи различных шкал между собой	309
7.13. Подход к определению операций над нечеткими значениями, задаваемыми функциями с несколькими максимумами	310
7.14. Модификация значений, вычисляемых нечетким алгоритмом	311
7.15. Анализ работоспособности резервированных систем на основе качественной информации, представленной нечеткими числами	315
7.16. Структура шкалы для измерения работоспособности блока	317
7.17. Анализ работоспособности блоков	317
7.18. Методика опроса экспертов для выяснения значений S_{ij} – степени причастности блока B_j к сбою M_i	320
7.19. Построение s_{ij}	320
7.20. Алгоритм построения s_{ij}	322
7.21. Прогноз интервала между сбоями	325
7.22. Алгоритм смены прогноза	326
7.23. Оценка правильности прогноза	327
7.24. Отключение процессов прогноза	328
7.25. Текущее среднее время	328
7.26. Определение работоспособности блока	329
7.27. Изменение значения качественной величины текущего среднего	330
Список литературы	331
Глава 8. Методы повышения эффективности генетических алгоритмов	333
8.1. Теория ГА, особенности и основные зависимости	333
8.1.1. Процесс оптимизации с помощью ГА и понятие схем	333
8.1.2. Теорема схем	339
8.1.2.1. Селекция	340
8.1.2.2. Кроссовер	341
8.1.2.3. Мутация	341
8.1.3. Повышение эффективности генетических алгоритмов	342

8.1.4. Разработка обобщенной схемы ГА	346
8.2. Исследование и разработка оператора селекции	349
8.3. Исследование операторов мутации и кроссовера	352
8.4. Разработка нечеткого адаптивного однородного кроссовера	353
8.4.1. Сущность и способы реализации	353
8.4.2. Разработка НБП для НАОК	355
8.4.3. Последовательный БГА+ с НАОК	358
8.4.4. Результаты экспериментов	363
Список литературы	368
Глава 9. Общая концепция и подход к построению распределенных самоорганизующихся информационных систем (РСИС)	370
9.1. Традиционные подходы к построению информационных систем	371
9.2. Классификация самоорганизующихся систем	373
9.3. Математическая модель самоорганизующейся информационной системы (СИС)	373
9.4. Структурно-функциональная модель СИС	375
9.5. Пример построения самоорганизующейся системы обмена информацией	380
9.5.1. Постановка задачи генерации правил в динамических сетях	382
9.5.2. Предварительная подготовка данных	383
9.5.3. Методика извлечения знаний из баз данных. Концепция нанесения сетки	384
9.5.4. Визуализация полученных данных	387
9.5.5. Применение нечеткой логики для прореживания данных в сетке	388
9.6. Параллельная реализация процесса построения самоорганизующейся системы обмена информацией	393
9.6.1. Многопоточность в РНР. Библиотека pthreads	396
9.6.2. Решение задачи предварительной очистки и трансформации данных без использования библиотеки pthreads	397
9.6.3. Параллельная реализация решения задачи предварительной очистки и трансформации данных с использованием библиотеки pthreads	399
Список литературы	403
Информация об авторах	407

ПРЕДИСЛОВИЕ

Уважаемые читатели

В этом году международная конференция по мягким вычислениям и измерениям SCM-2017 отмечает 20-летний юбилей. Поэтому данная монография посвящена этому событию

Коллектив авторов составили постоянные участники конференции, ведущие ученые нашей страны, научные труды которых развили направление мягких вычислений и мягких измерений и сформировали мощный научный базис современных методов и технологий обработки информации в условиях неопределенности. На основе трудов авторского коллектива сформированы новые научные школы по направлениям нечеткой математики, общей теории измерений, теории вероятностей и математической статистики, искусственного интеллекта, теорий управления и менеджмента и других.

Первый том монографии «Теоретические основы и методы» включает работы по созданию основных концепций, принципов, методологической базы мягких вычислений, мягких измерений, управления в условиях неопределенности, принятия управленческих решений.

Во втором томе отражены вопросы разработки алгоритмической основы, информационных технологий решения прикладных задач. Поскольку конференция носит междисциплинарный характер, прикладные задачи, рассматриваемые в монографии охватывают широкий круг сфер человеческой деятельности и служат ярким подтверждением эффективности и мощности и перспективности предлагаемых подходов и решений.

Материалы монографии позволяют оценить ретроспективы основных идей и школ, труды которых были представлены в томах, а также определить тенденции научных направлений и увидеть перспективы их развития.

Авторы благодарны организатору конференции – Санкт-Петербургскому государственному электротехническому университе-

ту в лице его ректората и членов Оргкомитета за предоставляемую им важную и полезную возможность обмена знаниями и обогащения перспективными идеями. Все эти годы тематика конференции не ограничивалась чисто техническими вопросами и сферами получения и преобразования научных знаний. Глубокие доклады философского и теологического плана были представлены на пленарных секциях докладами. Обсуждения таких работ неизменно проходили в живой дискуссии и много дали участникам для глубокого понимания сущности процессов формирования и применения принципов работы со знаниями, что, безусловно, способствовало расширению представлений о процессе формирования и преобразования знаний в живой природе и в системах и построениях искусственного интеллекта. Дискуссии ученых разных направлений знаний, несомненно, способствовали проявлению синергетического эффекта в каждой из областей научного творчества

Авторский коллектив надеется, что монография будет полезной как для состоявшихся ученых разных направлений науки, специалистов прикладных сфер деятельности, так и для молодых ученых, аспирантов и студентов, определяющих свой путь в науке.

Пожелаем нашей конференции и ее участникам многочисленных ярких событий, долгих творческих лет и самых светлых перспектив!

*Редактор монографии
проф. С.В. Прокопчина*

ГЛАВА 1

МЯГКИЕ ПОДХОДЫ К ИЗМЕРЕНИЯМ И УПРАВЛЕНИЮ СЛОЖНЫМИ СИСТЕМАМИ. РЕГУЛЯРИЗИРУЮЩИЙ БАЙЕСОВСКИЙ ПОДХОД

С.В. Прокопчина

1.1. Концептуальные основы решения задач в условиях неопределенности

1.1.1. Основные принципы решения задач измерения и управления сложными системами в условиях неопределенности

Информационная сфера в современном обществе стала важнейшим фактором его развития. Она непрерывно расширяет границы решаемых задач и технологических возможностей. Вместе с тем общество, и в первую очередь бизнес, ставят перед разработчиками информационных систем новые практические задачи, отличающиеся повышенной сложностью, междисциплинарностью используемых подходов, специфичностью и уникальностью объектов применения информационных технологий.

В отличие от предыдущих десятилетий повышаются требования к интеллектуальности информационных систем. Так, к числу основных требований к информационному обеспечению бизнес-процессов относятся требования по обеспечению получения решений в виде готовых рекомендаций, альтернатив, сценариев развития бизнес-процессов и бизнес-ситуаций.

В концептуальном плане реальные объекты и системы, с которыми взаимодействуют информационные технологии могут рассматриваться как сложные системы с позиций системного подхода.

Под сложными объектом или системой в работе понимается управляемая или частично управляемая система, состоящая из взаи-

мосвязанных природных, техногенных, экономико-социальных систем, активно взаимодействующая с внешней средой.

К числу таких систем относятся все технические объекты, процессы и комплексы, социально-экономические системы, энергетические, транспортные системы и объединяющие их территориальные системы, практические все объекты и системы, которые связаны с деятельностью человека от управления корпорацией до управления регионами и странами.

Сложность измерения свойств таких систем и управления ими общеизвестна. Это объясняется спецификой самих систем, их уникальностью, невозможностью повторения событий, процессов и ситуаций, отсутствием средств измерений их свойств, часто имеющих качественный характер, множеством неизученных и меняющихся влияющих факторов и связей между ними, что в целом создает условия неопределенности, состоящие в неполноте, неточности, нечеткости информации. В таких условиях не «работают» классические методы и модели измерений и управления.

Это относится, прежде всего, к методам измерений в их классической детерминированной постановке, методам классической математической статистики, методам эконометрики.

На практике приходится сталкиваться с подгонкой адекватных моделей свойств и деятельности реально функционирующих систем к ограничениям, предположениям, допущениям и постулатам предлагаемых математических методов и моделей, что значительно сужает сферу их применения, а в ряде ситуаций неприменимо. Это обусловлено прежде всего тем, что адекватные модели построить в условиях неопределенности достаточно сложно. Кроме того, активное влияние на систему факторов окружающей среды определяет изменчивость системы, что приводит к необходимости перестройки модели в ходе функционирования реальной системы и обработки информации. Для построения адекватных моделей необходимо накапливать и перерабатывать все имеющиеся об объекте измерения или управления данные и знания.

Такая же ситуация сложилась и с прикладными компьютерными информационными системами, используемыми в настоящее время для решения измерительных и управленческих задач. Их функционирование базируется на программном обеспечении, основанном на «жестких» моделях и алгоритмах по сути с заранее заданным набором или ограниченным пространством решений. Кроме того, пользова-

тель, решающий подобные задачи, должен приспосабливаться к условиям реализации этих алгоритмов на компьютерах, зачастую в совсем неудобном и неорганичном для прикладной задачи режиме представления данных, знаний и получаемых решений

В современных измерительных и управленческих науках формируются новые подходы к решению задач в условиях неопределенности. Однако, эти разработки до настоящего времени не систематизированы с точки зрения общей методологии создания систем измерения и управления, функционирующих в условиях значительной неопределенности.

Все сказанное определяет актуальность создания методологии разработки технологий обработки данных и решения задач в условиях неопределенности.

Целесообразно выделить принципы, которые являются основными для построения гибких и «мягких» подходов и технологий к решению вышеуказанных задач, эффективных в условиях значительной неопределенности.

Первый принцип. Решение измерительных и управленческих задач связано с моделированием, оценкой характеристик и параметров, идентификацией ситуаций и процессов. В условиях неопределенности решение этих задач характеризуется неполнотой, неточностью, нечеткостью исходной информации; нечеткостью постановки задачи, требований, ограничений, критериев; неустойчивостью, нечеткостью, неоптимальностью получаемых решений. Часто оптимальное решение для таких условий вообще не существует. Это дает возможность сделать вывод о невыполнении требований Адамара, а именно, существования, единственности и устойчивости решений. Таким образом, информационные задачи в условиях неопределенности являются некорректными и должны ставиться и решаться как обратные задачи восстановления значений параметров, вида функциональных зависимостей, генерации выводов по экспериментальной информации. Таким образом можно утверждать, что в условиях неопределенности задачи мониторинга, аудита, нормирования, управления, планирования, являются некорректными по Тихонову и могут быть введены в класс квазиустойчивых применением регуляризации. Для успешной реализации вычислительного процесса решения прикладной задачи в условиях неопределенности постановка задачи из-

мерения или управления должна быть определена в виде совокупности прямых и обратных задач восстановления значений, состояний, ситуаций по полученной информации. При получении новой информации об объекте прикладной задачи процесс решения задачи на основе методов оптимизации с поиском экстремума целевой функции повторяется в итерационном режиме. При этом в условиях неопределенности экстремум целевой функции может быть определен в вероятностной, условной форме и представлен в виде списка условных альтернативных значений с соответствующими оценками, (в частности, значениями) степени их достоверности. В общем случае и сама целевая функция может быть представлена в нечеткой форме, уточняемой в процессе накопления информации при решении задачи. Это принципиальное замечание связано с необходимостью в условиях неопределенности ее снижения за счет поступления новых знаний, то есть изучения объекта прикладной задачи и окружающей его среды. В такой постановке технологии решения обратных задач и в целом прикладных задач измерения и управления могут быть определены как интеллектуальные технологии решения обратных задач и интеллектуальные технологии решения задач измерения и управления

Второй принцип. Для обеспечения устойчивости решения прикладной задачи в условиях неопределенности необходима регуляризация решений на компакте решений.

Это связано с необходимостью обрабатывать информацию из информационных потоков со значительной степенью энтропии, добавляющей неопределенность в информационную ситуацию, делающей получаемые решения размытыми, нечеткими, неполными и неточными. Для устранения и недопущения неконтролируемого добавления энтропии необходимо обеспечить непрерывность измерительно-вычислительно-аналитического процесса получения решения прикладной задачи. То есть необходимо, чтобы компакт решений, объединяющий компакты всех частных решений был непрерывным и целостным. Это может быть обеспечено только на основе непрерывного контроля качества входных информационных потоков и получаемых решений, их сопряженности и согласованности друг с другом согласно последовательности этапов процесса обработки информации.

Выбор метода регуляризации определяет возможности, мощность и сферу применимости технологии или алгоритма.

При отсутствии достаточной информации ограничения, формирующие пространство компакта решений могут быть определены в нечеткой или вероятностной формах. Это обуславливает формирование компакта решений также в нечеткой или вероятностной формах. Ограничения могут изменяться во времени, что ведет к динамичности границ компакта решений. Все в целом формирует нечеткий динамический компакт решений.

Нечеткий динамический, мягкий компакт решений D – это компакт с нечеткими динамическими границами, состоящий из совокупности возможных компактов D_i ($i=1, I$), с соответствующими им степенями их вероятности (P_i , $i=1, I$) или возможности.

Формализуя вышесказанное можно записать:

$$D = * D_i | P_i, \quad (1)$$

где «*» – символ свертки компактов при создании единого информационного пространства решений.

Третий принцип. Развиваемость функциональных и континентных возможностей методов, моделей, алгоритмов.

Это – вопросы обеспечения динамичности компактов решений, функциональных алгоритмических наборов, моделей, измерительных шкал и, ограничений и нормативов, что обусловлено меняющимися свойствами прикладных систем и объектов.

Для реализации гибких моделей и методов измерения и управления в условиях неопределенности необходимо, чтобы пространство решений, модели, измерительные шкалы, измерительные и управленческие технологии могли изменяться во времени в соответствии с изменением свойств, структур и функций объектов измерений и управления.

Динамичность мягкого компактов, в частности, может быть отражена формулой:

$$D(t) = *_{i=1, I} D_i(t) | P_i(t), \quad (2)$$

где $t=1, T$, и

t – момент времени формирования мягкого компакта;

T – период функционирования компакта.

Четвертый принцип. Метрологическое обеспечение моделей, алгоритмов, решений. Оно должно сопровождаться параллельным основному процессу обработки информации процессом вычисления комплексов метрологических характеристик, получаемых промежуточных и основных решений. Такие комплексы должны включать показатели точности, надежности и достоверности решений с вычислением риска их практических применений, энтропии и количества информации, получаемой на данном этапе процесса обработки информации. В задачи метрологического сопровождения и обеспечения должны входить как задачи метрологического анализа и вычислений показателей качества получаемых решений, так и задачи метрологического синтеза для получения алгоритмов обработки с требуемыми метрологическими характеристиками и для обеспечения управляемости информационно-технологическими процессами. **Инии** метрологического обеспечения с обратными связями для корректировки показаны на рисунке 1.

Пятый принцип. Информационные технологии обработки информации в условиях неопределенности должны иметь обратные связи с возможностью коррекции алгоритмов, моделей, решений с целью их непрерывного контроля и улучшения, что иллюстрирует рисунок 1.

Шестой принцип. Когнитивность и интерпретируемость решений не только для понимания их, но и для активного использования и вовлечения в процесс обработки информации знаний, опыта, интуиции и логики экспертов, лиц, принимающих решения (ЛПР) и пользователей как активных интеллектуальных компонентов систем, моделей и алгоритмов обработки информации. При создании когнитивных систем и сетей эти активные компоненты могут рассматриваться как интеллектуальные агенты сетей. В когнитивных измерительных сетях эти компоненты рассматриваются как виртуальные измерительные приборы.

Седьмой принцип. Возможность обработки и свертки как числовой, так и качественной, лингвистической информации для полноты использования как данных так и знаний в процессе обработки, интеграции и передачи решений на высшие уровни иерархии моделей.

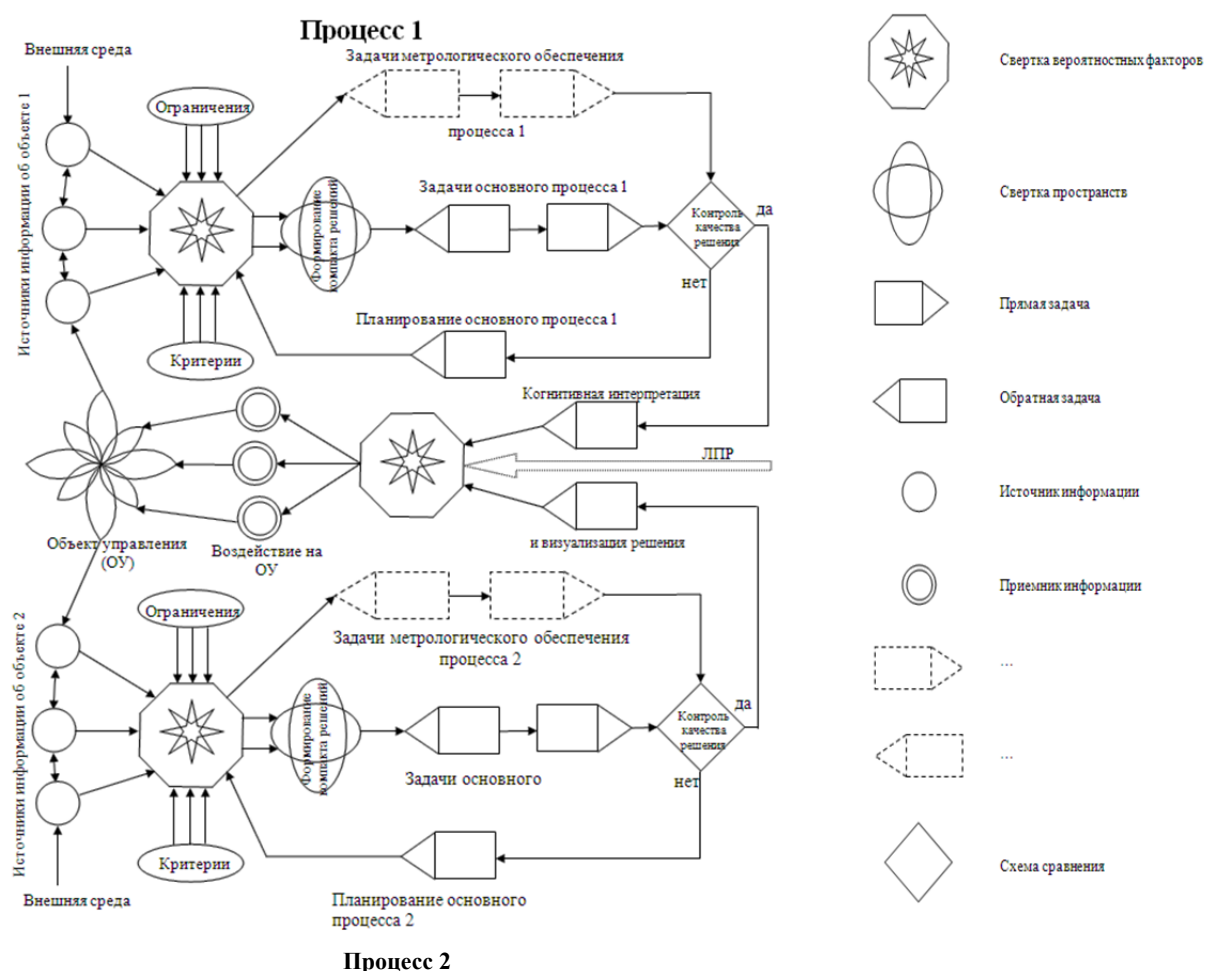


Рис. 1. Обобщенная технология реализации основных принципов решения задач измерения и управления сложными объектами в условиях неопределенности

Восьмой принцип. Итерационность и последовательная взаимосвязанность процессов сбора и интеграции информации, идентификации моделей объектов и оценки их параметров (решение обратных задач), а затем определение искомых решений основной целевой задачи (решение прямых задач) и вновь повторение этого цикла с новыми данными и знаниями, поступившими во время предыдущего цикла.

Девятый принцип. Условность полученных решений (в виде моделей, рекомендаций, оценок, выводов, сценариев развития событий) в рамках принятых допущений, ограничений и требований. В условиях неопределенности, как уже отмечалось, необходимо решение обратных задач измерения, оценивания, идентификации. При решении этих задач в такой постановке сами ограничения и до-

пущения являются меняющимися и в значительной мере неопределенными с точки зрения их абсолютной истинности. Это обуславливает их условный и вероятностный характер, что приводит к необходимости вычисления степени вероятности, возможности, достоверности получаемых решений. Таким образом, при изменении влияющих факторов самого объекта или окружающей среды, необходимо корректировать постановку задачи, формулируя новые требования, критерии (нормативы) и ограничения, адекватные новым условиям.

Очевидно, что перечень принципов не полон, но существенен для создания технологий решения измерительных и управленческих задач в условиях значительной неопределенности.

Можно утверждать, что, если в теоретических построениях или практических методиках не учтен хотя бы один из вышесформулированных принципов, трудно ожидать полезного решения прикладных задач в реальных условиях их реализации. Условия неопределенности сопровождают решение практически всех антропогенных задач.

1.1.2. Обобщенная методика и схема решения задач в условиях неопределенности

Изменение свойств сложного объекта и взаимосвязей объекта и внешней среды определяет необходимость постоянной коррекции модельных представлений, оценок свойств, критериев и других компонентов информационного пространства. В условиях значительной неопределенности модель сложного объекта и среды его окружения должна изменяться в зависимости от получаемой информации и меняющихся требований, ограничений, целевых функций и критериев решаемой задачи. В концептуальной форме это можно определить как изменение степени «погружения» модели $G^{(M)}(t)$ в систему объекта $G^{(O)}(t)$ и формально представить в виде выражения (3):

$$G^{(O)}(t) \rightarrow G^{(M)}(t), \quad (3)$$

где $G^{(O)}(t)$ – система динамического объекта $G^{(O)}(t) = Q^{(O)}(t) * R^{(O)}(t)$ со свойствами $Q^{(O)}(t)$, отношениями $R^{(O)}(t)$, меняющимися в зависимости от времени t ;

$G^{(M)}(t)$ – система динамической модели $G^{(M)}(t) = Q^{(M)}(t) * R^{(M)}(t) * L^{(M)}(t)$ со свойствами $Q^{(M)}(t)$, отношениями $R^{(M)}(t)$ и ограничениями,

допущениями, требованиями $L^{(M)}(t)$ постановки задачи, также меняющимися во времени [56].

Важной задачей является создание методологических и информационно-технологических разработок, обеспечивающих более полное и глубокое отражение системы объекта в модели на основе развития структуры модели.

Информационные технологии должны позволять, получая новую информацию об объекте и среде его функционирования, использовать ее для перестройки информационного пространства модели в соответствии с содержанием новой информации, требованиями, ограничениями и критериями для определенной прикладной задачи.

Обобщенное уравнения решения измерительных и управленческих задач в условиях неопределенности в оптимизационной форме можно записать в следующем виде:

$$\{h_k(Q,t)|\{MX\}_k(Q,t)\} |(G^{(o)}(Q,t)|P^{(G0)}|(Y^{(o)}|(O,t)|P^{(YO)})*(G^{(oe)}(Q^{(OE)}(t))|Y^{(oe)}(t)|P^{(YOE)}) = \{\arg \text{extr } C(t)|P^{(C)} [f_j(\{x_i(t)\})|(G^{(o)}(Q,t)|P^{(G0)}|(Y^{(o)}|(O,t)|P^{(YO)})*(G^{(oe)}(Q^{(OE)}(t))|Y^{(oe)}(t)|P^{(YOE)})]\} \quad (4)$$

где – $\{h_k(Q,t)|\{MX\}_k(Q,t)\}$ – нечеткое, мягкое решение задачи с комплексом характеристик его качества (точности, надежности, достоверности и др.) в условиях $Y(G^{(o)}(Q,t)|P^{(G0)}|(Y^{(o)}|(O,t)|P^{(YO)})*(G^{(oe)}(Q^{(OE)}(t))|Y^{(oe)}(t)|P^{(YOE)})$, определяемых (условными с вероятностью P или степенью доверия P) априорной информацией, требованиями, ограничениями, допущениями, моделями объекта прикладной задачи и окружающей его среды, функционального базиса $F(t)$ для обработки информационных потоков $X_I(t)$; в условиях влияющих факторов окружающей среды $\{MX\}_k(Q)$ – комплекс метрологических характеристик качества решений;

$\{C_z(t)\}$ – критерий, задающий логику решения задачи;

$F(j)$ – технологии, задающие и реализующие регуляризацию решений;

$\{x_i(t)\}$ – совокупность потоков информации, используемой для получения решений.

Первым этапом решения любой информационной задачи о реальном объекте является этап его предварительного моделирования.

Естественным требованием к модельным представлениям динамических и эволюционирующих объектов являются требование адекватного отражения всех свойств объектов и взаимодействий с окружающей средой, исторических этапов и тенденций их развития. Однако, в реальных условиях решения указанных задач представляется весьма трудным выполнить данное требование из-за принципиальной невозможности абсолютно полного познания таких объектов. Информация о них всегда неполна, неточна, нечетка и объективно или субъективно искажена, что не позволяет построить полностью адекватную модель объекта априори. Скорее можно сделать предположения и допущения о характере и классе модели. Таким образом, априорная модель формируется часто не столько на объективных частотных данных, сколько на некоторых знаниях о возможности применения той или иной модели для конкретных типов и условий решения задач.

В зависимости от типа ограничений и допущений, принятых при решении конкретной задачи (в том числе и измерительной), жесткости требований к качеству ее решения, и полноты и обоснованности априорной информации информационные ситуации можно условно разделить на три основных типа.

Первый тип информационной ситуации определим как тип с незначительной априорной неопределенностью информации об объекте измерения (моделирования, контроля, оценивания, управления) и среде его функционирования. Ко второму типу можно отнести ситуации с ограниченной априорной неопределенностью, при которой принимается допущение, что адекватная модель объекта принадлежит некоторому заранее определенному классу моделей с параметрической неопределенностью. Этот тип ситуаций соответствует классу адаптивных измерений, рассмотренному в ряде работ 70-х – 80-х годов прошлого столетия. Третий тип информационных ситуаций соответствует ситуациям с неограниченной (значительной) априорной неопределенностью, связанной со спецификой сложных объектов, в частности, с принципиальной невозможностью полного познания свойств самого объекта и его взаимодействия со средой, а также с его непрогнозируемой изменчивостью в процессе функционирования.

Первому типу соответствует достаточно обширный и жесткий круг ограничений и допущений, значительно упрощающей ее до степени параметрической неопределенности. Модель при этом считается полностью адекватной объекту и известной с точностью до искомого

параметра. К числу таких задач относятся все задачи простых прямых и косвенных измерений физических величин, имеющих вполне определенное соответствие между принятой системой их эталонов и реальных воплощений.

Классические методы измерений значений физических величин всегда ориентированы на этот тип информационных ситуаций. Шкалы измерений такого рода обычно являются метрическими, с сильными вычислительными свойствами и очень ограниченной семантикой. Они выбираются заранее и не меняются в процессе решения задачи. Результатом решения является вполне четко определенное числовое значение определяемого параметра. Однако, в тех задачах, о которых шла речь выше, искомым результатом является не конкретное числовое значение, а скорее степень интенсивности проявления в окружающей среде изучаемого свойства, знание о силе его связи и влиянии на компоненты среды. Этап подобной интерпретации числового значения всегда в этих задачах перекладывается на субъекта-измерителя. Таким образом, при наличии значительной априорной неопределенности неконтролируемые субъективные искажения получаемого результата вносятся уже на первом этапе измерительного процесса при субъективном определении модели объекта (из-за неадекватности модели объекту), а затем на этапе его интерпретации при достижении цели решения.

Аналогичная ситуация происходит и в задачах управления в информационной ситуации первого типа. Управление производится по жесткому алгоритму и заранее выбранной модели. по сути по входной информации определяются только неизвестные параметры модели управления.

Линейность и простота алгоритмов измерений и управления в классической постановке определена жесткой однозначно определенной линией стратегии реализации измерений и управления. Поэтому план и модели измерительного эксперимента или процесса управления четко определены до его начала и не меняются в ходе решения задачи. В первой информационной ситуации они являются единственно возможными и не требуют пересмотра и корректировки. Таким образом, при первой информационной ситуации неопределенности в плане и моделях в информационных технологиях нет.

Как было отмечено выше, вторая информационная ситуация характеризуется ограниченной непараметрической неопределенностью

и соответствует адаптивным алгоритмам организации процессов получения и обработки данных. Стратегия планирования реализации информационных технологий предполагает его циклическую разветвленную, но вполне четко определенную структуру и поэтому, хотя в процессе реализации информационных технологий измерения и управления и требуется этап оптимизации моделей и параметров, но при практической верности допущений, определяющих условия эксперимента как ситуацию второго типа, неопределенности для всего класса моделей, как и в первой ситуации не существует.

Однако, несмотря на некоторую вариабельность, возможность выбора модели объекта, как правило, в этих стратегиях их состав определяется содержанием того класса моделей, внутри которого ищутся решения задач. Для развивающихся объектов и меняющихся условиях решений указанных выше задач такой жесткий круг модельных представлений безусловно неприемлем. Основным свойством информационных технологий в ситуации третьего типа (в подобных ситуациях действительно реализуются не алгоритмы, а именно информационные технологии с включением субъективного фактора их реализации) должна быть их способность к реализации непрерывного познания объекта. Это свойство при работе со сложными объектами в третьей информационной ситуации выходит на первый план. Только непрерывно получая знания об объекте и влияющих на него факторах, а затем используя их в процессе получения решения, можно достичь цели поставленной задачи с приемлемой для практики точностью. Таким образом, класс моделей объектов и влияющих факторов внешней среды должны быть легко модифицируемыми в ходе решения задачи в соответствии с поступающими и получаемыми данными и знаниями. На основании этого модельные представления должны корректироваться на основе новых знаний, модель объекта должна развиваться в согласии с развитием объекта, аккумулируя в себе знания обо всех новых проявлениях объекта. В соответствии с этим должна развиваться и видоизменяться система шкал, представляющих новые свойства объекта и среды, а также их взаимовлияния. Их жесткая определенность, характерная для информационных технологий, ориентированных на второй тип информационных ситуаций, не позволяет реализовать такие требования, что и обуславливает необходимость разработки гибких стратегий реализации измерительного эксперимента и процесса решения задачи, меняющихся при

смене условий эксперимента и поэтому предполагающих оптимизацию его плана и параметров в ходе его осуществления. В условиях значительной неопределенности здесь могут быть внесены дополнительные неопределенности в процесс решения задачи, которые трансформируются в неопределенности результатов. Заметим, что при наличии статистической априорной информации для коррекции этих неопределенностей можно использовать вероятностный подход к представлению и описанию погрешностей и неопределенностей на этом этапе, но такой опыт имеется далеко не всегда. Чаще всего, как отмечалось выше, имеются лишь предположения и интуиция экспериментатора. В таком случае необходимо рассматривать эту информацию как гипотетическую и использовать для описания неопределенности субъективные вероятности или возможностный подход. Это обуславливает необходимость привлечения специальных математических методов, позволяющих включить субъективный опыт в структуру информационных технологий.

На основании сказанного можно сформулировать общую схему решения задач в условиях неопределенности, иллюстрируемую рисунком 2 и представить ее в виде последовательности этапов.

Основные этапы организации процесса решения задач в условиях неопределенности:

1. Постановка задачи в нечеткой формулировке, определение цели задачи в многоальтернативном виде.
2. Идентификация информационной ситуации и типа задачи. Если ситуация относится к третьему типу (значительной неопределенности), то перейти к пункту 3.
3. Формирования пространства входной информации, мягких компактов исходных информационных потоков, ограничений, допущений, критериев.
4. Метрологическое обоснование источников информации.
5. Формирование моделей объекта измерения или управления на основе решения обратных задач измерения, оценивания и идентификации, а также моделей окружающей среды в нечеткой форме, в форме списка возможных моделей объекта измерения или управления с определением их достоверности.
6. Выбор репрезентативных показателей свойств объектов согласно разработанных моделей и их свертка в единое информационное пространство.

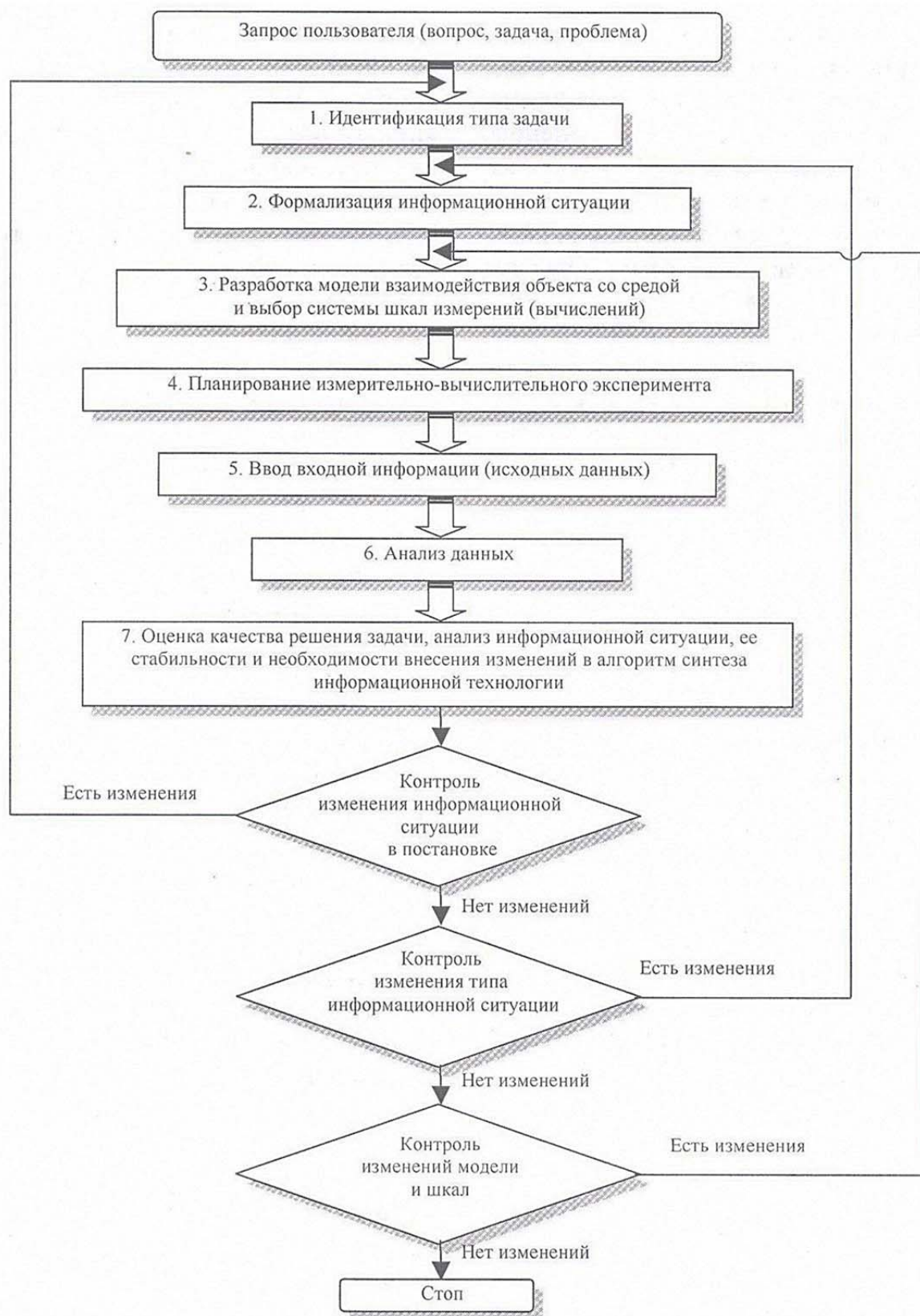


Рис. 2. Обобщенная схема реализации информационных технологий измерений и управления в условиях неопределенности

7. Выбор правила вывода, критериев и логики принятия решений.
8. Формирование и метризация пространства мягких компактов решений.
9. Создание конкретных измерительных шкал для свойств объекта и внешней среды.
10. Решение обратных задач измерения, оценивания, идентификации.
11. Планирование и организация процесса решения основной задачи с целью обеспечения требуемого качества решений.
12. Реализация процесса решения основной задачи с получения решений в виде набора альтернатив с определением их возможности, надежности, достоверности, риска, информативности и т.д.
13. Проверка качества решений.
14. Коррекция всех составляющих процесса определения решений для следующего момента времени.
15. Повторение процесса, начиная с первого этапа для получения уточненного решения задачи в динамическом режиме.

Отметим, что схема получения решений в условиях неопределенности разработана в соответствии со стандартом ISO 9001.

Основу ее построения составляет принцип непрерывной перестройки ее структуры в целях повышения эффективности и качества решений. Важно, что для реализации этого принципа все составляющие блоки схемы, включая и блоки постановки задачи, уточнения целей задачи, определения модели и идентификации информационной ситуации охвачены обратной связью для коррекции их и непрерывного развития в процессе получения решения.

1.1.3. Методологические основы регуляризирующего байесовского подхода для решения задач в условиях неопределенности

Одним из подходов, позволяющих решить вышеуказанные задачи в новой постановке, соответствующей условиям неопределенности, учитывая перечисленные принципы, является регуляризирующий байесовский подход (РБП).

РБП был создан автором в восьмидесятых годах прошлого столетия [1–7]. В это же время были разработаны основные технологии математической обработки данных и технологии решения прикладных задач на основе РБП.

Являясь модификацией байесовского подхода (вероятностного вывода на основе известной формулы Байеса), РБП ориентирован на

условия значительной неопределенности, при которых могут быть неизвестны составляющие байесовской формулы.

Определение составляющих байесовской формулы производится в ходе вычислительного эксперимента. Для их определения в процессе вычислений используются все данные и знания, представленные как в числовой, так и в лингвистической формах. Поэтому информационные технологии решения прикладных задач на основе РБП, названы байесовскими интеллектуальными технологиями (БИТ).

На основе прикладных БИТ решаются задачи изучения свойств сложных объектов, классификации, измерений, мониторинга, аудита, нормирования, управления в условиях неопределенности, то есть малых выборок, уникальности данных и ситуаций, многообразия информационных потоков, нечеткости и неполноты информации.

В работах [1–17] рассмотрены вопросы теоретического обоснования и практического использования РБП для решения широкого круга прикладных задач и показана реализация вышеуказанных принципов в байесовских интеллектуальных технологиях (БИТ).

Рассмотрим кратко идеи, реализованные в РБП и БИТ, отражающие вышесформулированные принципы.

Основу РБП создали три известных классических подхода: системный, измерительный и байесовский подходы.

Первый – системный подход используется в РБП не только для изучения объектов и формирования ряда репрезентативных свойств объекта, но и для системного синтеза, состоящего в объединении, интеграции этих факторов в единую условную концептуальную модель объекта и внешней среды. Системный синтез обеспечивает условную адекватность (квазиадекватность) модели объекту и среде его функционирования.

Измерительный подход создает условия для получения устойчивых, надежных и решений обеспечивает создание условно-устойчивых (квазиустойчивых) измерительных моделей, путем регуляризации решений на шкалах специального типа, названными шкалами с динамическими ограничениями (ШДО). Конструкции ШДО подробно разработаны для различных типов данных и задач и подробно рассмотрены [3–15], а также будут приведены ниже.

Байесовский подход используется для формирования и реализации правил логического вывода и получения решений прикладных задач. Основным критерием байесовского подхода, как известно, яв-

ляется средний риск получаемого решения. Байесовское правило оптимизирует процесс получения решения по данному критерию и обеспечивает минимизацию среднего риска решения.

Задачи определения состояний и свойств сложных объектов на основе РБП состоят в непрерывном изучении свойств и характеристик этих объектов путем обобщения прошлого опыта работы с ним и вновь поступающей информации с позиций измерительного подхода, принципиальную основу которого составляет метрологическое обоснование получаемых решений. Новые знания, получаемые в ходе такого измерительного процесса, объединяются с архивами на основе байесовской свертки и служат априорной информацией для будущих экспериментов. Причем, чем объемнее и разнообразнее поступающая информация, обобщаемая далее на основе принципов теории измерений и метрологии, тем полнее и достовернее получаемые результаты.

Идейная основа РБП, заложенная в классическом байесовском подходе, его индуктивная (обобщающая) логика послужила основой для реализации принципа интеграции и свертки информации.

В выводах индуктивной логики байесовского подхода возможно получение решений с определенной степенью сомнения, выражаемой количественной мерой апостериорной байесовской вероятности (достоверности) решения. Эта величина вычисляется по модифицированной формуле Байеса для РБП.

Постулатами РБП служат принципы классического байесовского подхода:

- свойства объектов, внешней среды, их характеристики и параметры считаются неопределенными;
- результаты наблюдений и экспериментов считаются неслучайными событиями, порождающими ряд гипотез о послуживших их появлению причинах: событиях, процессах, ситуациях;
- априорная информация объединяется с поступающей на основе интегральной свертки информационных потоков для получения апостериорного распределения гипотез;
- вывод или решение принимается на основе оптимизационного правила, минимизирующего риск принятия полученного решения или максимизирующего полезность и безопасность решения.

Эти принципы, развитые позднее в работах ученых-байесианцев, позволили прийти к понятию субъективной (фидуциальной) вероят-

ности, что создало научную основу для привлечения неформализованного опыта специалистов к принятию решений. Эта возможность в свою очередь обусловила возможность эффективной работы с выборками малого объема, так как, по выражению Д. Эванса, «знания могут компенсировать отсутствие данных».

Априорная неопределенность исходной информации и основных компонент байесовского правила, имеющая место в большинстве практических прикладных задач, может быть определена в измерительных задачах как измерительная ситуация в условиях неопределенности. В этих условиях задачу оптимального байесовского оценивания можно рассматривать как задачу восстановления измеряемой характеристики по экспериментальной неполной информации, что делает задачу байесовского оценивания в математической постановке некорректной, а получаемые решения неустойчивыми. Для получения однозначного и устойчивого решения задача байесовского оценивания должна быть введена в класс корректных, что определяет необходимость модификации подхода в направлении придания ему регуляризирующих свойств. В РБП такая регуляризация реализуется на основе многомерного и многоальтернативного шкалирования.

Одним из способов повышения степени устойчивости решений является ограничение пространства возможных решений путем формирования системы дополнительных ограничений, приводящих к созданию условного (в рамках поставленных ограничений и допущений) условно полного компактного пространства устойчивых решений.

Основная идея регуляризирующего байесовского подхода (РБП) состоит в привлечении измерительного подхода в процесс байесовского оценивания путем применения шкалирования, метризирующего и регуляризирующего пространство решений. Все модификации классического байесовского подхода являются некорректными по Тихонову обратными задачами. Поэтому в условиях неопределенности их решения являются неустойчивыми и неединственными. Получение результатов на компактах иерархических (согласно структурам объектов) метрических шкал, согласно теореме Тихонова [28], обеспечивает их устойчивость в пределах этих компактов. Таким образом, введение измерительной шкалы, обуславливает дополнительные ограничения на пространство возможных решений, оказывает регуляризирующее действия, что обосновывает название данной моди-

фикации байесовского подхода и позволяет ввести задачу измерений или управления на основе РБП в класс условно-корректных по Тихонову. Алгоритмы РБП, реализующие такую идейную основу будут обладать регуляризирующими свойствами при проведении измерений или управления в сложных измерительных ситуациях при значительной априорной неопределенности информации об свойствах и структуре объекта измерения или управления и влияющих факторах внешней среды.

Отметим также важные свойства РБП подхода:

- подход обладает свойством внутреннего единства методологии, что делает его универсальным для применений в параметрических, функциональных и системных моделях и позволяет строить соответствующие модели иерархической структуры на основе единой методологии, обеспечивающей формализованную связь всех видов ИИ, системность и комплексность их результатов;

- подход максимизирует объективность и полноту решений, так как дает возможность учета всей совокупности данных и знаний при получении решения;

- принципы подхода обеспечивает последовательное накопление и коррекцию знаний об объекте изучения, что позволяет реализовать основную функцию мягких подходов и технологий (познания объекта) и создает основу для развития модельных представлений ОИ и ВФ;

- при использовании подхода в процессе длительного изучения объекта, характерного для задач в условиях неопределенности, имеется принципиальная возможность получения все более точных результатов так как по мере накопления знаний байесовские решения концентрируются вокруг истинного решения (заметим, что данное положение вполне согласуется с принципом теории измерений о возможности постоянного увеличения точности измерений [43]);

- на основе байесовского подхода имеется возможность перспективного и ретроспективного изучения свойств объекта на основе индуктивной логики, что особенно важно для получения объективных и достоверных прогнозов и восстановления истории развития объекта;

- в связи предыдущими свойствами подход может быть эффективным для реализации его средствами новых технологий на основе средств получения и обобщения знаний, в частности, в виде экспертных систем и интеллектуальных комплексов.

Для построения алгоритмов РБП необходимо определить их принципиальную основу, которая характеризуется оптимизационным уравнением для получения эффективного и достоверного решения.

В обобщенной форме уравнение БИИ может быть представлено в следующем виде [15]:

$$\{h_k | P_k\} = \left\{ \begin{array}{l} \operatorname{argmin} C_B[\varphi_j(x_i | \{Y_i\})] \\ h_i \in H_k \\ x_i \in X_i \\ \varphi_j \in \Phi_j \end{array} \right\}, \quad (5)$$

где h_k – результат БИИ в форме значения параметра, аналитического выражения функции, вывода или рекомендации, являющийся элементом множества решений H_k , представляющего собой носитель соответствующей шкалы БИИ;

P_k – вероятность появления результата при данных условиях БИИ УГ, поступившей в ходе БИИ экспериментальной информации $X_i \in X$, реализованном алгоритме φ_j из множества алгоритмов Φ_j ;

C_B – байесовское решающее правило, оптимизирующее выбор решений из множества.

Байесовская апостериорная вероятность результата определяется по формуле Байеса, принимающей на множестве H_k вид:

Рассчитанное по (5) значение апостериорной вероятности $P(h_k(t))$ решения определяет его достоверность на компакте решений H_k при условиях измерений, определяемых $Y(t)$. Из уравнения (5) следует, что решение задачи на основе РБП представляется списком альтернатив с соответствующей им степенью вероятности или возможности их осуществления, то есть в нечеткой форме, определяемой апостериорным законом распределения вероятностей $F(h(t))$ возможных решений на множестве H_k .

Реализация познавательной функции измерения или управления при активном взаимодействии объекта измерения или управления с развивающейся внешней средой, ее влияющими факторами (ВФ), при меняющихся условиях измерений, связана с необходимостью адекватного изменения моделей объекта и ВФ, что обуславливает требование синтеза алгоритмов РБП в процессе измерений или управления. Рассмотрим основные вопросы, связанные с синтезом алгоритмов РБП.

1.1.4. Модели и шкалы с динамическими ограничениями. Концепция, типы, структура

Концепция модели с динамическими ограничениями (МДО) в рамках РБП предполагает построение иерархической динамической модели объекта и среды в процессе управления системой сложного объекта или поведением объекта на основе непрерывного познания свойств этих систем и допускает возможность изменения как состава компонентов моделей, отражающих свойств объекта среды и формирующих МДО и их взаимоотношений, так и самих систем представления свойств и отношений.

Числовой информации в задачах и системах в условиях неопределенности, как правило, бывает недостаточно для развития моделей объектов и среды в соответствии с их меняющимися свойствами. Поэтому целесообразно использовать дополнительную информацию, представленную в различных формах, в том числе и лингвистической информации, что безусловно повышает качество решений, обогащая семантическое содержание результатов, снижая тем самым степень их неопределенности при интерпретации. Однако, неметрические шкалы, предназначенные для отображения лингвистической информации, обладая значительной семантической наполненностью (предельным вариантом является семантика естественных языков), не обладают мощными вычислительными аспектами метрических шкал, необходимыми для обеспечения скорости, объективности, функциональной широты допустимых вычислительных преобразований.

Очевидно, соединение, «сопряжение» тех и других типов шкал дало бы возможность значительно повысить эффективность методов и средств измерений в направлении их интеллектуализации, так как, с одной стороны, этот приближало бы их к семантической наполненности и гибкости естественных языков, с другой стороны, обеспечивало бы возможность быстрой и надежной переработки значительных объемов данных и знаний.

Концепция байесовских интеллектуальных измерений (БИИ), основанная на РБП, представляет собой методологию синтеза нового типа моделей и шкал для реализации «обобщенных» измерений с целью достижения качественного решения прикладной задачи на основе всестороннего познания свойств СО и среды его функционирования. Такие модели и шкалы [1–5] реализуются в метрических пространствах динамических компактов их носителей и носят название *шкал с*

динамическими ограничениями (ШДО), что обусловлено их способностью к адаптивному изменению своей структуры в процессе, накопления информации о развитии свойств объектов или среды их функционирования. Теоретические основы построения ШДО, синтеза их иерархических структур и информационные технологии их применения в прикладных областях приведены в работах автора [1–17]. Процесс решения прикладной задачи на основе методологии РБП реализуется как процесс целенаправленного преобразования иерархической структуры моделей и шкал (ШДО), способных адекватно отражать свойства эволюционирующего объекта. Иерархическая структура моделей и шкал (ШДО) отражает свойства объекта и среды, геометрически представляя собой некий гиперкуб условно связанных факторов объекта и среды, а для шкалированного представления моделей гиперкуб взаимосвязанных одномерных ШДО.

Результаты решений представляются в качестве реперов шкал с меняющимися свойствами согласно поступающей информации. В совокупности своей эти взаимоувязанные шкалы представляют собой гипершкалу (гиперкуб), которая реализует модель объекта и среды его функционирования в виде МДО, которая может быть достигнута путем введения шкал.

Шкала БИТ может менять свои свойства и структуру (носитель, реперы, состав допустимых отношений и др.) согласно изменению структуры МДО. Поэтому она названа шкалой с динамическими ограничениями (ШДО). Ее вид приведен на **рисунке 3**.

Шкала типа ШДО для измерения свойств одномерного показателя является двумерной шкалой, по одной из осей которой откладываются значения показателя в числовой или лингвистической формах, по другой – степень достоверности (уверенности, возможности) результата. При добавлении числа контролируемых показателей в многомерном пространстве параметров строится многомерная шкала, которая при переходе на новый, более высокий уровень иерархии сворачивается в интегральный показатель, для которого формируется новая двумерная шкала. Таким образом осуществляется развитие структуры ШДО согласно структуре МДО. Согласно этому возникают новые ветви информационных технологий измерения новых показателей, их контроля и аудита, интерпретации ситуаций, генерации рекомендаций и так далее. Таким образом осуществляется непрерывное развитие моделей и информационных технологий на основе РБП.

Глава 1. Мягкие подходы к измерениям и управлению сложными системами. Регуляризирующий байесовский подход

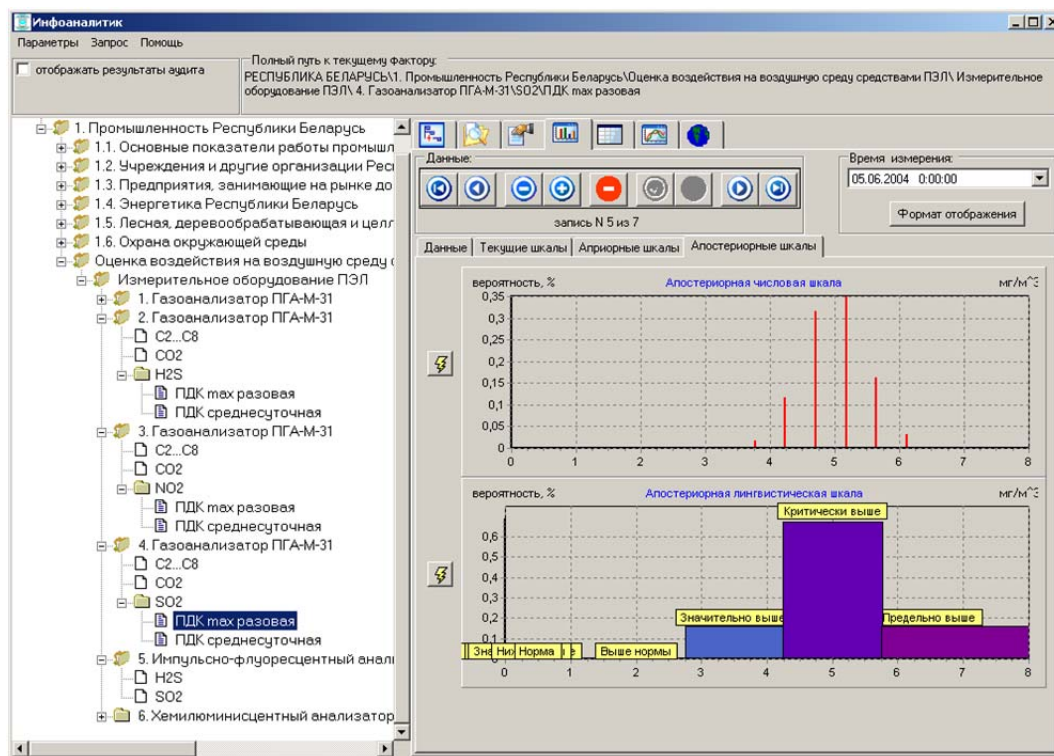


Рис. 3.

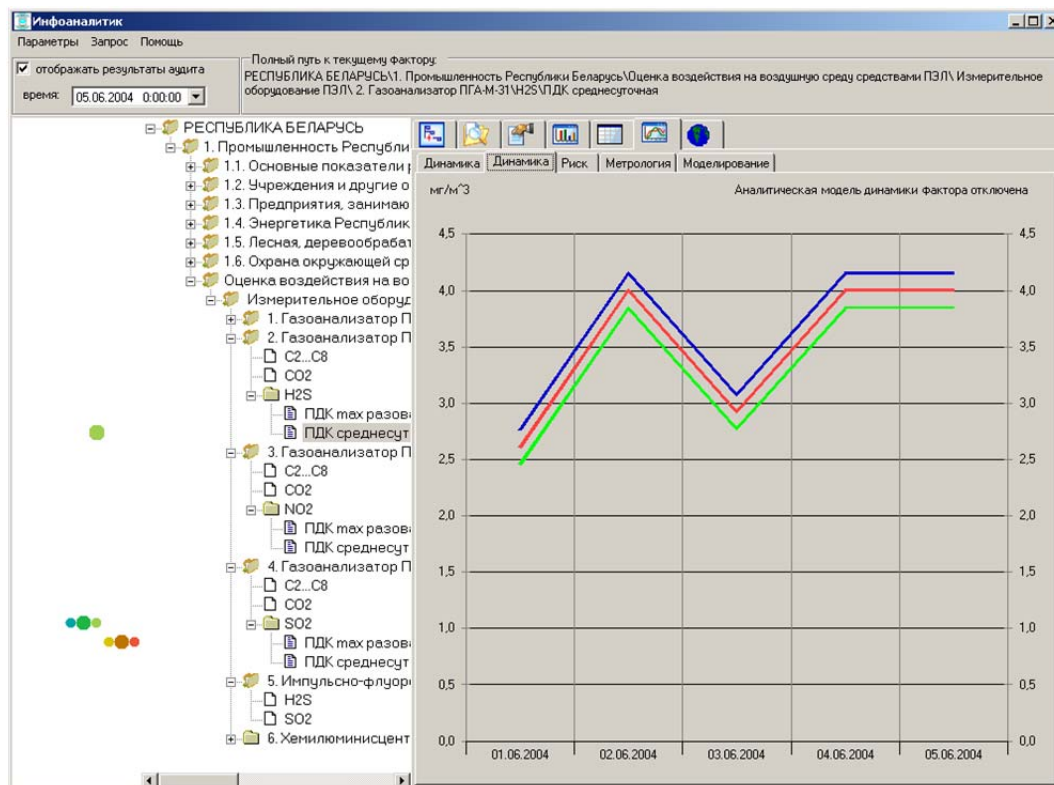


Рис. 4.

В основе синтеза информационных технологий РБП применяется *принцип единства измерений*, позволяющий согласовать входы и выходы отдельных шкал и преобразовывать их в соответствии с функциональным содержанием информационной технологии и соблюдением метрологических требований к решениям информационной системы. Для этого параллельно с вычислительным процессом реализуется процесс метрологического сопровождения каждого решения в виде показателей точности, надежности, достоверности, энтропии и риска. Эти показатели объединяются в комплексы метрологических характеристик. **Рисунок 3** иллюстрирует вид интерфейса пользователя, отражающего перечисленные показатели для динамической модели контролируемого свойства. Сама динамическая БИТ-модель приведена на **рисунке 4**.

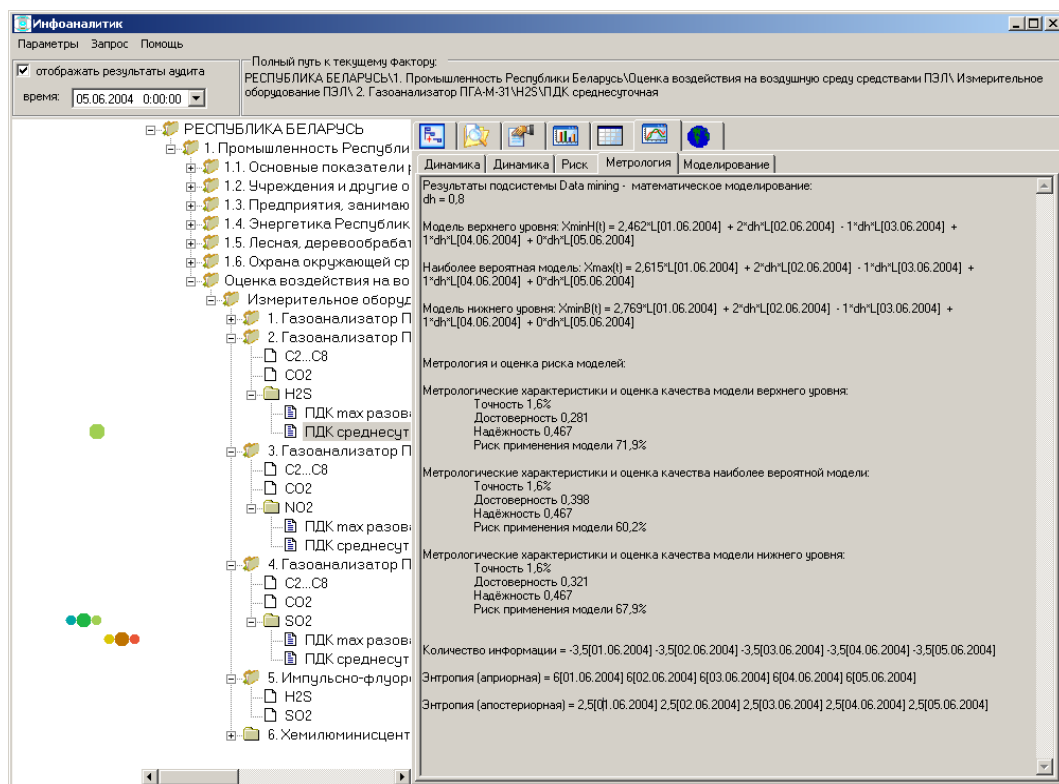


Рис. 5.

Решение, полученное на основе байесовских интеллектуальных технологий, (БИТ-решение) представляет собой ряд альтернативных оценок свойства с соответствующими комплексами метрологических характеристик и представляет собой регуляризованную байесовскую оценку (РБО), свойства которой рассмотрены в [1].

Уместно заметить, что в условиях значительной неопределенности каждое элементарное решение из состава РБО не обладает высокими показателями надежности и достоверности. Однако в целом РБО покрывает истинное значение с достаточно высокими показателями качества и минимумом риска. Вид РБО для различных информационных ситуаций представлен на рисунках 3, 6, 9 и иллюстрирует решения при различном количестве информации или различной энтропии.

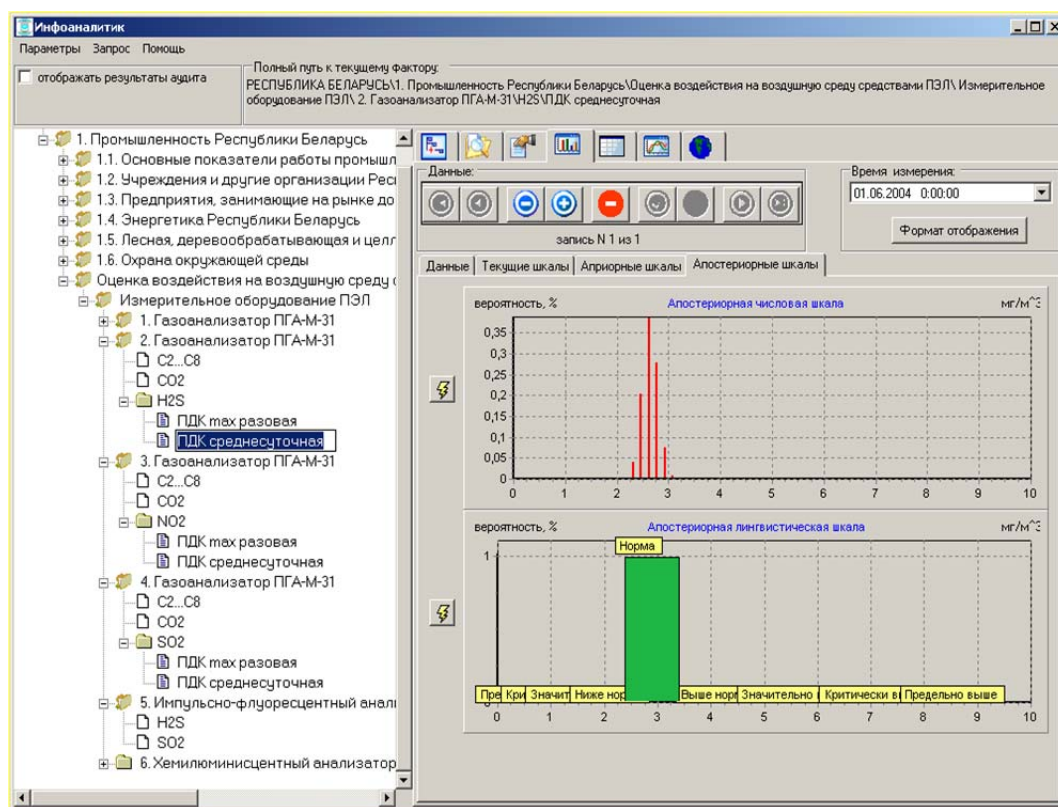


Рис. 6.

По сути РБО в целом представляет собой дискретный закон распределения измеряемого свойства. Для получения аналитического вида закона распределения РБО. применяются технологии функциональных байесовских интеллектуальных измерений, рассмотренные ниже в параграфе 1.2.

Рисунок 5 иллюстрирует функциональные возможности байесовских интеллектуальных технологий по построению и преобразованию ШДО.

На рисунке 6 приведены результаты аппроксимации результатов, полученных на ШДО в виде апостериорных распределений РБО ти-

повыми законами распределений из классов системы Пирсона, для чего используется математическая библиотека РБП.

Рис. 7.

Полученные модели законов распределений используются затем для интерпретации оцениваемых свойств объекта, для генерации выборок большего объема и последующей ее обработки мощными методами математической статистики, для управления качеством получаемых решений и их риском.

Существует несколько типов ШДО.

Самыми простыми с точки зрения системы отношений и начальными шкалами являются классификационные ШДО, на которых задается структура модели типа МДО в виде иерархической структуры ее взаимосвязанных показателей. Такая шкала имеет вид дерева факторов или целей. На ней реализуются отношения эквивалентности. По классификации общей теории репрезентативных измерений такая шкала соответствует шкале наименований. Она имеет самое богатое семантическое содержание и не имеет возможности реализации ка-

ких-либо преобразовательных функций. В информационных системах на основе БИТ пользователь самостоятельно может легко менять структуру классификационной ШДО, постоянно развивая ее. Возможно изменение структуры такой ШДО автоматически с приходом новой порции данных или знаний.

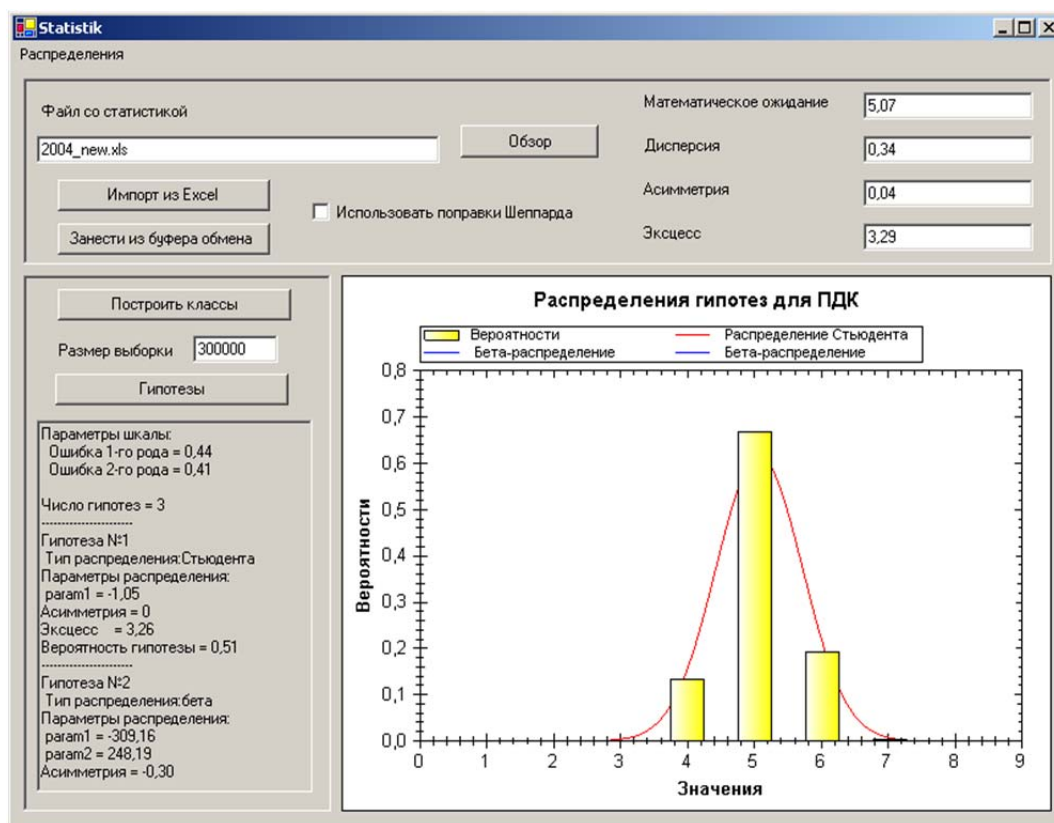


Рис. 8.

Для оценки параметров и показателей применяются статические и динамические параметрические оценочные шкалы, вид которых представлен на рис. 1, 2, 4. Для динамических показателей реализованы динамические ШДО (см. рис. 2). Причем, как первый, так и второй тип шкал реализован виде сопряженных ШДО для обработки и хранения как числовой, так и лингвистической информации. Числовые ШДО имеют тип шкал отношений и обладают богатыми вычислительными возможностями. В зависимости от изменения характеристик информационных потоков меняется структура этих шкал.

Для функциональных преобразований созданы функциональные ШДО, теория построения которых приведена в [1].

Задачи нормирования, критериального выбора и аудита производятся на критериальных шкалах. При смене критериальной базы или изменении свойств контролируемого объекта, когда меняется норма показателей его свойств, шкала настраивается на новые значения норм автоматически или пользователем. Таким образом может контролироваться динамика норм и критериев. Кроме того, эти шкалы могут быть сопряженными для реализации контроля одного и того же показателя по множеству различных критериев.

Критериальные шкалы могут быть полезны в задачах контроля (экологического, энергетического, технического и других) для установления предупреждающих границ при вероятностном представлении контролируемого показателя. Например, существует практически важная задача оценки вероятности превышения значений контролируемого параметра допустимого предела и в соответствии с этим изменения предупреждающих границ контроля. На основе БИТ на функциональных ШДО в режиме мониторинга определяется изменение закона распределения параметра (см. рис. 7), например, в следствии изменения технологического оборудования, режимов производства, материала заготовок. По результатам оценки закона распределения параметра принимаются решения об изменении предупреждающего значения на критериальной шкале БИТ, что предупреждает появление брака, экологических или энергетических рисков. Подробнее механизм реализации этого подхода рассмотрен в [1].

Для интерпретации ситуации и генерации управленческих рекомендаций используются динамические ситуационные шкалы типа ШДО, структура которых также может перестраиваться в зависимости от изменения содержимого баз знаний и данных.

ШДО БИИ для определения состояния или степени интенсивности проявления контролируемого свойства может быть представлена в виде трехзвенной структуры ШДО БИИ, приведенной на рис 8, представляющем собой экранную форму системы поддержки принятия управленческих решений (СППУЭР), реализующей методологию БИИ.

Данная ШДО состоит из априорной, текущей и апостериорной ШДО и используется для свертки информационных потоков, разнообразных по форме представления информации, ее содержанию и степени неопределенности.

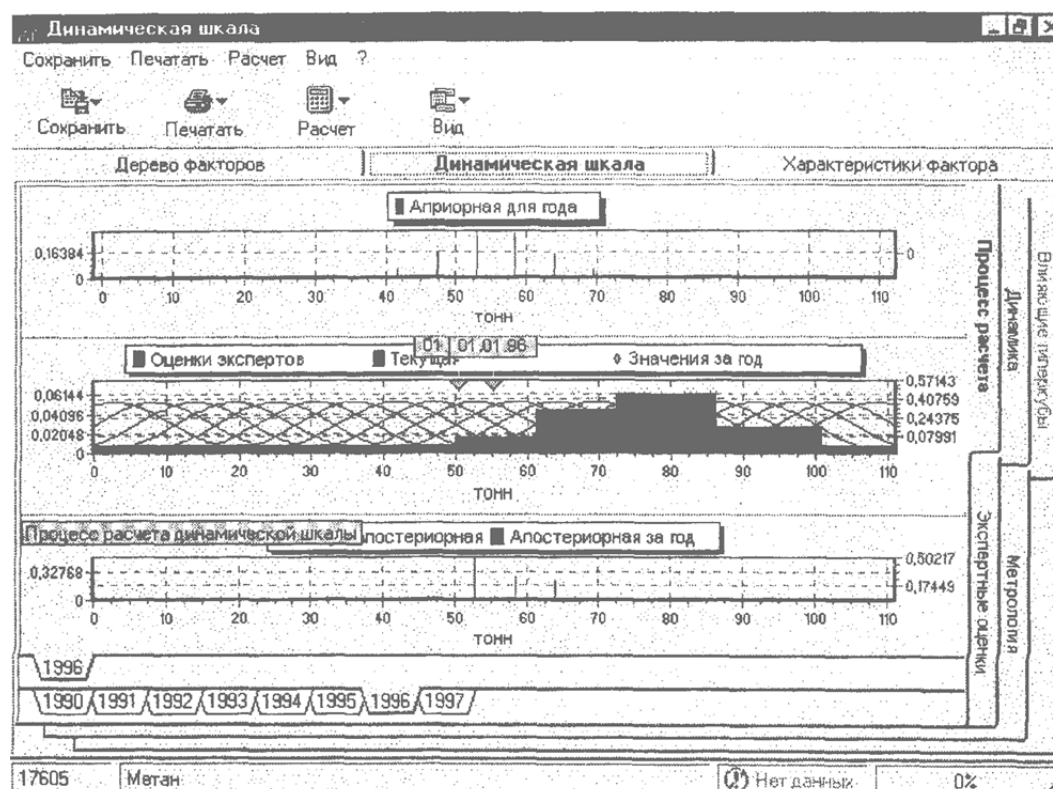


Рис. 9. Трехзвенная ШДО для измерения на основе РБП

Таким образом средствами информационных технологий РБО при использовании ШДО достигается:

- интеграция разнообразной по форме представления информации с целью повышения или достижения требуемого качества результата;
- метрологическое обоснование полученных решений в виде количественных показателей меры апостериорной (остаточной) неопределенности, например, показателей точности, надежности, достоверности;
- реализация принципа саморазвития моделей объектов измерения и среды их функционирования на основе адаптации структур ШДО к свойствам СО, познаваемым в процессе БИИ.

ШДО БИИ имеет в своей структуре сопряженную лингвистическую шкалу, пространство трехзвенного носителя которой представляет собой иерархическое развитие известной структуры пространства нечеткого терм-множества лингвистической переменной [1, 3–5], для которой базовым множеством X является носитель текущей шкалы БИИ контролируемого свойства СО. При этом впервые возникает возможность учета лингвистической информации на основе формализмов РБП в процессе получения результата измерений, что позво-

ляет использовать дополнительную лингвистическую информацию для снижения степени неопределенности и повышения качества результата измерения. Кроме того, подобная трехзвенная структура лингвистической ШДО позволяет интегрировать на основе методологии БИИ априорную и текущую лингвистическую информацию, получая апостериорное лингвистическое решение с метрологическим обоснованием его качества, а также идентифицировать, или адаптировать форму функций принадлежности, объективизируя ее путем привлечения дополнительной информации (и числовой) в процессе измерений или обработки. В работе [1] приведены теоретические положения синтеза сопряженных лингвистических ШДО.

1.1.5. Основные принципы интеграции информации на основе регуляризирующего байесовского подхода (РБП) и его технологий

Интегрирующие свойства байесовского подхода являются принципиальной основой для разработки информационных технологий свертки информации, ее преобразований на многомодельной основе, генерации гибких стратегий управления, комплексной оценки ситуаций и разработки сценариев их оптимального развития. Основным препятствием к широкому использованию этого подхода до настоящего времени являлась неустойчивость и чувствительность байесовских оптимальных решений при неточной, неполной, нечеткой информации.

Методологическая основа интеграции разнотипной, неполной, неточной и нечеткой информации на основе РБП представлена в работах [1, 2, 3]. Данная модификация байесовского подхода направлена на получение устойчивых оптимальных байесовских решений в условиях значительной неопределенности информации. Информационные технологии интеграции на основе РБП строятся на основе шкал с динамическими ограничениями, которые, кроме основных шкал для представления информации (числовой или лингвистической), имеют сопряженные универсальные шкалы для передачи информации на следующие этапы ее обработки.

Универсальные ШДО имеют типизированную структуру носителя и допустимого множества отношений для каждого типа решения, стандартную форму представления информации, позволяющую абстрагироваться от конкретных единиц представления информации. Они представляют собой порядковые шкалы с допустимыми отношениями эквивалентности и порядка и стандартными наборами реперов

в виде терм-множеств (обычно 9 элементов для обеспечения приемлемой практической точности решений). К числу таких шкал относятся несколько видов типовых шкал для оценки состояний объектов, значений параметров технических систем и степени интенсивности проявления свойств природных компонентов окружающей среды; интерпретации, динамики и тенденций развития ситуаций, а также генерации рекомендаций по управлению ситуациями и сценариями их реализации с учетом всех значимо влияющих факторов внешней среды.

Значения параметров и характеристик свойств сложных объектов обычно определяются со значительной степенью неопределенности, что обуславливает целесообразность привлечения дополнительной информации об объектах и их свойствах. Интеграция такого типа данных может производиться на *параметрических шкалах ШДО* по технологиям БИТ в соответствии с формой представления и степенью неопределенности информации. Данные технических измерений с использованием приборов и информационных систем при этом могут быть на формальной основе методологии РБП проинтегрированы с экспертными оценками и расчетными значениями. Так, при имеющейся совокупности потоков данных $\{X_i\}$, $i=1, \dots, I$ на основании БИТ-технологий можно сформировать интегральный поток данных $X_I^{(o)}$:

$$X_I^{(o)} | Q_I^{(o)} = \bigstar_{i=1}^I X_i | Q_i = X_{I-1}^{(o)} | Q_{I-1}^{(o)} \star X_I | Q_I;$$

В качестве информационных потоков данных рассматриваются: результаты измерений средствами информационно-измерительных систем; данные статистической отчетности, анкетирования или учета; результаты моделирования, расчетные данные, результаты функциональных преобразований (трансформированные экспериментальные данные) измерительной или вычислительной информации; физические константы; картографическая информация, видео-, аудио-информация; экспертная информация; фактологическая информация.

Степень неопределенности потоков данных, априорные знания о них и ограничения представляются совокупностью условий интеграции потоков.

Свертка информации производится по модифицированной формуле Байеса:

$$P(h_{kj}^{(o)} | X_I^{(o)} | Q_I^{(o)}) = \frac{P(h_{kj} | x_{ij} | X_i | Q_i) \cdot P(h_{k,i+1,j} | x_{i+1,j} | X_{i+1} | Q_{i+1})}{\sum_{l=1}^K P(h_{lij} | x_i | X_i | Q_i) \cdot P(h_{l,i+1,j} | x_{i+1} | X_{i+1} | Q_{i+1})}$$

где $h^{(o)}$ – интегральная (для совокупности потоков информации $X|^{(o)}$) регуляризированная байесовская оценка (РБО); -РБО для X , потока данных.

Метрологическое обоснование каждого типа информационных потоков X ; ведется в соответствии с разработанными формализованной методической базой метрологии БИИ и принципами преобразования комплексов метрологических характеристик [1, 3], которые включают показатели точности, надежности (уровней ошибок 1 и 2 рода) и байесовской достоверности для каждой РБО. Свертка комплексов метрологических характеристик MX выполняется в соответствии со следующим выражением:

$$\{MX\}_{kj} = \underset{i=1}{*}^I \{MX\}_{kij} | X_I^{(o)} | Q_I^{(o)};$$

Закон распределения общей погрешности определения интегральной числовой оценки значения физической величины для БИТ-технологий определяется как композиция законов распределения погрешностей определения РБО для отдельных потоков данных в виде их свертки:

$$f(\Delta h_{kj}) = \underset{i=1}{*}^I f(\Delta h_{ij}) | X_I^{(o)} | Q_I^{(o)};$$

При использовании БИТ-технологий в автоматическом режиме решается задача определения многомерного закона распределения в аналитическом виде, сложность которой общеизвестна. В рамках реализации БИТ-технологий свертки законов распределений разработана технология аппроксимации апостериорного распределения кривыми из системы Пирсона. Круг аппроксимирующих распределений может расширяться за счет включения новых семейств распределений. При идентификации распределения используются вся имеющаяся априор-

ная информация при полном метрологическом контроле получаемых результатов, что позволяет при заданных требованиях к качеству аппроксимации снизить необходимый для этого объем выборки.

Для интеграции лингвистических РБО в БИТ-технологиях применяется модифицированная формула Байеса для свертки функций принадлежности

Интеграции вероятностной и возможностной информации производится на основании методологии РБП и БИИ в виде сопряженной регуляризированной байесовской оценки в соответствии с формулой:

$$\begin{aligned} \mu(h_{kj}^{(o\lambda)} | h_{kj}^{(o)} | X_i^{(o)} | Q_i^{(o)}) = \\ = \frac{\mu(h_{kj}^{(\lambda)} | h_{kj} | x_{ij} | X_i | Q_i) \circ \mu(h_{k,j+1,j}^{(\lambda)} | h_{k,j+1,j} | x_{i+1,j} | X_{i+1} | Q_{i+1})}{\sum_{l=1}^K \mu(h_{lj}^{(\lambda)} | h_{lj} | x_i | X_i | Q_i) \circ \mu(h_{l,j+1,j}^{(\lambda)} | h_{l,j+1,j} | x_{i+1} | X_{i+1} | Q_{i+1})} \\ h^{(ou)} = h_{kj}^{(o\lambda)} * h_{kj}^{(o)}, \text{ где} \end{aligned}$$

Методологическая основа свертки такого рода рассмотрена в [1].

Уместно отметить, что для практических задач смена логики вывода решения может играть важную роль при индексации ситуации по степени ее значимости. Примеры рассмотренные в [4], позволяют рекомендовать использование логики Заде и соответствующего аппарата вычислений при необходимости тщательного, детального рассмотрения всех возможных сценариев развития событий и ситуаций, а «жесткие» логики (типа логики Лукашевича) могут быть рекомендованы для решения задач в условиях значительной неопределенности, в которых необходимо получить единственное (или «слабо» размытое), наиболее целесообразное решение.

Сопряженная шкала числовых значений и лингвистических оценок состояний объектов наблюдения – статическая ШДО, представлена на **рисунке 7**, который иллюстрирует определение интегральной оценки значения концентрации примеси метана, полученной на основании интеграции расчетных данных и априорных экспертных оценок (на рисунке определена как «текущая» ШДО). Оценка получена на числовой шкале как нечеткое значение концентрации примеси метана, а на интегральной апостериорной лингвистической шкале как нечеткая оценка степени проявления этого компонента воздушной среды при транспортировке ресурсов газа

Носитель *интегральной универсальной лингвистической шкалы* представляют реперы экологического нормирования («норма», «ниже

нормы», «выше нормы») содержания примеси метана в воздушной среде, что (при использовании этого типа шкалы для оценки степени проявления других примесей) позволяет, абстрагируясь от единиц измерения компонентов воздушной среды, получить интегральную оценку состояния воздушной среды по совокупности рассматриваемых примесей.

Применение интегрирующих технологии РБП в таких условиях особенно актуально, так как интегрирующие свойства их необходимы для восстановления пропущенных данных, уточнения решений, составление практически приемлемых прогнозов, ретроспективных оценок и сценариев.

1.1.6. Метрологическое обоснование информационных технологий решения задач в условиях неопределенности

Метрологическое обоснование информационных технологий решения задач в условиях неопределенности позволяет производить оценку качества информации каждого источника данных и каждого получаемого решения в виде комплексов метрологических показателей: точности, надежности, достоверности, риска, энтропии, количества информации

Точность определяется по формуле (6):

$$\xi_s = \frac{\max_{h_s \in H_k} \rho(h_s, h_{s+1})}{\rho(h_K, h_1)}, \quad (6)$$

где $\rho(h_K, h_1)$ – диапазон шкалы H_K ,

$\max_{h_s \in H_k} \rho(h_s, h_{s+1})$ – максимальное расстояние между соседними элементами носителя шкалы;

Надежность. Надежность результата характеризует устойчивость решения. Показатель надежности основан на уровнях ошибки первого и второго рода и определяется как (7):

$$V_s = (1 - \alpha_s)(1 - \beta_s), \quad (7)$$

где α_s – уровень ошибок первого рода (отражающий вероятность отвержения правильного решения на шкале);

β_s – уровень ошибок второго рода (характеризующий вероятность принятия неправильного решения на шкале).

Говоря об ошибках первого и второго рода можно провести аналогию с проблемой вероятности принятия неверных решений при приемке товаров. В этом случае α (риск производителя) – это вероятность отказа от приемки партии продукции, которая оказалась бы приемлемой с точки зрения потребителя. А β (риск потребителя) – это вероятность приемки партии продукции, неприемлемой для потребителя.

Уровень ошибок первого и второго рода на шкале выбирается как максимальный из всех вероятностей ошибок первого и второго рода гипотез шкалы.

Достоверность. Показателем достоверности каждой гипотезы на шкале служит апостериорная вероятность его появления, определяемая по формуле Байеса. Достоверность самой шкалы при этом принимается сумме достоверности гипотез и соответственно равна единице. Однако во многих случаях из практических соображений с целью удаления малозначащих гипотез на шкале оставляют только набор гипотез, удовлетворяющих некоторому критерию значимости. В таком случае достоверность шкалы определяется как (8):

$$R = R^a \cdot \sum_{h_j \in H_r} P(h_j) \quad (8)$$

где R – окончательная достоверность решения на шкале;

R^a – достоверность шкалы до удаления не значимых гипотез;

H_r – множество значимых гипотез на шкале.

Риск – величина, показывающая риск принятия данного решения. Вычисляется, как:

$$(1 - R),$$

где R – достоверность.

Это позволяет создать единую технологию с возможностью управления качеством получаемых решений.

Важным методологической задачей является создание цепочек информационной ценности для оптимизации структуры вычислительного процесса в условиях значительной неопределенности. Вычисление метрологических комплексов для решений позволяет определять для каждого этапа вычислительного процесса степень энтропии и количество информации, получаемой на этом участке процесса обработки информации. Для оптимизации вычислительного процесса и повышения его эффективности можно оптимизировать его этапы в соответствии с их информационной ценностью, определяемой количественной мерой информации, получаемой на них. Таким образом, в прикладных задачах можно отбрасывать те этапы, на которых либо не добавляется полезной информации, либо прибавляется ее слишком мало. Следуя этому принципу, можно оптимизировать блоки структур систем управления, например управляющих подразделений корпораций, структур департаментов министерств и администраций или других управленческих систем.

1.2. Организация измерительных процессов в условиях неопределенности. Регуляризирующий байесовский подход

1.2.1. Исторические аспекты развития теории измерений и интеллектуализации измерительных процессов

Начиная с конца восьмидесятых годов под влиянием успехов методологий искусственного интеллекта и необходимости решения указанных выше задач в условиях неполноты и недостоверности информации, начало формироваться новое направление в теории измерений: *интеллектуальные измерения*. К тому времени уже стали понятны и ощутимы жесткие ограничения шествующих методов измерений, а поиски новых методологических платформ для реализации познавательной функции измерений естественным образом вели к методам искусственного интеллекта, объединяющего под своей эгидой самые разнообразные способы и средства получения знаний.

Взаимопроникновение методов искусственного интеллекта, а также мягких вычислений, на всех этапах своего развития обращавшихся к теории меры – основы всех измерительных процессов – и измерительного подхода особенно ясно прослеживается в настоящее время. В современных работах по теории и практике измерений, ней-

ротехнологиям, нечетким и гибридным системам ведутся разработки способов и средств взаимодействия методологий этих научных направлений, реализации новых классов информационных технологий и систем получения и обработки информации в условиях неопределенности. Здесь и далее обобщения автора основаны на обзорах современных научных работ, библиография которых приведена в статьях и монографиях.

Теоретическую базу таким разработкам создали фундаментальные работы по общей теории измерений, теории меры, теории нечетких множеств, теории возможностей и их прикладные методологии. Эту тенденцию подтвердили многочисленные доклады, сделанные в 90-х годах на конференциях по мягким вычислениям, особенно таких, как конференции международной ассоциации нечетких систем IFSA, конференции и труды Международного комитета по измерениям ИМЕКО. Особое внимание при этом уделялось вопросам определения мер и способов контроля неопределенности данных и знаний, получаемых в процессе решения различного рода прикладных задач, а так же контролю качества и метрологии решений. В настоящее время теория измерений активно развивается в направлении изучения семантических свойств измерительных шкал и их эпистемологии.

Направление «интеллектуальные измерения» было заявлено в 1986 году на УШ Конгрессе ИМЕКО и сразу же получило активное развитие. Так, X Конгресс ИМЕКО практически уже был полностью посвящен этой и смежными с ней тематикам. Примерно в это время были впервые опубликованы работы по применению методов *интервальной арифметики*, *теории нечетких множеств*, методов обработки лингвистической информации в измерительных процессах. Появились работы по применению средств экспертных систем в измерительной практике.

Вопросам изучения семантических, интерпретационных свойств измерений и измерительных шкал посвящено немало интересных научных работ. В связи с этим необходимо упомянуть работы Л. Финкелыптейна, В. Розенберга, Е. Вошни, Д. Хоффмана, В. Кнорринга, А. Фьока, С. Муравьева, К. Мари, А. Фаллоу, Е. Паппа и др.

Фундаментальные работы Ю. Пытьева и других ученых по вовлечению в измерительный процесс нечисловой информации и методологии обработки лингвистической информации, теории нечетких множеств, теории возможностей дали основу автору для создания но-

вого вида измерений – *нечетких, «возможностных» измерений*. В таких измерениях участвуют информационные потоки данных и знаний, полученных не на детерминированной или вероятностно-статистической основе, а на основе определения меры гипотетической возможности их появления. В качестве нечеткого результата выступают множества числовых или лингвистических значений или выражений с соответствующим показателем меры возможности каждого из элементов таких множеств. В рамках этого направления появились работы по привлечению этих методологий для метрологического обоснования результатов измерений.

Появление этого научного направления было связано со значительным количеством новых измерительных задач, в которых аппарат традиционных, классических измерений характеристик физических величин оказывается малоэффективным. С одной стороны такие задачи характеризуются высокой степенью неопределенности всех составляющих измерительного процесса: неточностью, недостоверностью и неполнотой экспериментальной информации об объекте измерения, условиях измерений, системе измерения, и даже цели задачи; их изменчивостью во времени, сложностью структур объектов измерения, динамичностью характеристик и сложностью взаимосвязей, не позволяющих решить задачу на основе алгоритмов прямых или косвенных измерений. С другой стороны, как правило, в таких задачах требуется не высокая точность измерений, а скорее примерная оценка, но с точностью, достаточной для решения прикладной задачи. Кроме того, обычно имеется достаточно много дополнительной информации в виде данных и знаний, прямо или косвенно отражающие свойства объекта измерения и среды его функционирования, которая должна быть эффективно использована, но не учитывается в моделях измерительных процессов классической теории измерений. Причем, значительная часть такой информации представлена в виде текстов, графиков, аналитических зависимостей и других видов нечисловой информации. Практически все экономико-хозяйственные, управленческие, природопользовательские, природоохранные, социальные, политические и другие реальные задачи относятся к данной категории задач.

На практике такие задачи сопровождаются сложными условиями проведения экспериментов, связанных с наличием значительной априорной неопределенности о свойствах объектов и влияющих факторах среды их функционирования, взаимосвязях между ними, неточ-

ностью и неполнотой экспериментальной информации, недоступностью для непосредственных наблюдений многих свойств объектов или влияющих факторов, что выделяет познавательную функцию методологии их решения как основополагающую. Поэтому постановка указанных задач как измерительных обуславливает усиление роли познавательной функции измерений и выдвигает требование получения результатов их решений в форме знаний (аналитических выражений для моделей, а также выводов и решений) на основании учета всего объема априорной и поступающей в процессе измерительного эксперимента информации, в том числе и нечисловой.

Выполнение этого требования способствовало привлечению аппарата теорий принятия решений, искусственного интеллекта, нечетких множеств к созданию новых информационно-измерительных технологий для обработки разнородных потоков экспериментальных данных средствами измерительной техники.

Научные работы в этом направлении привели к созданию теории и систем интеллектуальных измерений в начале 90-х годов. Основной акцент научной и практической новизны был сделан на измерительной обработке знаний, метрологического обоснования знаний, использовании наряду с числовыми данными лингвистических переменных [,,,].

В указанных работах использован регуляризирующий байесовский подход, и результат интеллектуальных измерений представлен в виде множества значений с соответствующими комплексами метрологических характеристик. «Нечеткий» результат измерений позволял «мягко», с определенной долей точности и риска, достаточных для решения прикладной задачи, оценивать свойства объекта измерения.

Особую роль в формировании теории и приложений интеллектуальных измерений сыграло введение в структуру информационно-измерительной системы (ИИС) процессорных средств, предоставивших широкие технические возможности для вычислительного усложнения измерительных алгоритмов, реализации технологий обработки и получения знаний, принятия решений и выводов. Работы отечественных и зарубежных ученых создали теоретическую основу для появления поколения измерительных средств, реализующих алгоритмы нового типа. Они обладают функциональными возможностями, позволяющими реализовать принципиально новый тип измерительного процесса на базе сложных информационных технологий, включа-

ющих на каждом этапе измерений комплекс измерительных и метрологических преобразований. Это позволяет организовать единую измерительную цепь преобразований результатов первичных измерений, априорной информации об объекте измерения и среде его функционирования в конечное решение и рассматривать эту цепь как цельный измерительный процесс, алгоритмическая основа которого охвачена схемой метрологической поверки с реализацией принципов единства измерений

Рассмотрим некоторые концептуальные аспекты интеллектуальных измерений (ИИ).

Измерение – это процесс сопоставления свойств одного объекта (сравнительного эталона, образцового объекта) Q с аналогичными свойствами другого (объекта измерений) Q_x на определенном метрическом пространстве S в соответствии с метрикой M и мерой L , определенными в соответствии со свойствами данного метрического пространства и измеряемыми свойствами объекта измерений.

При реализации измерительного процесса логика сопоставления может принимать мягкую и жесткую формы. В зависимости от принятой логики результат измерений может формироваться либо в виде четкого числового значения, либо в виде набора альтернативных значений числового или лингвистического типа. По сути этим и определяется главное отличие ИИ от традиционных измерений.

Выбор логики определяется прежде всего целью и условиями решения измерительной задачи. Со сменой цели может быть изменена и логика измерений.

Для реализации мягкой логики в измерительном процессе используется аппарат теории принятия решений, главным принципом которого в условиях неопределенности является формирование системы альтернативных решений. Применительно к измерительным задачам это означает формирование результата измерений в виде набора альтернативных значений, составляющих в совокупности «нечеткий» результат. Каждая составляющая этой совокупности должна иметь свое метрологическое обоснование в виде комплекса метрологических характеристик, характеризующих качество данного результата. В ряд таких характеристик могут входить показатели точности, надежности, достоверности, информативности, риска и других метрологических характеристик.

Именно метрологическое обоснование каждой измерительной альтернативы, осуществляемое параллельно с реализацией процесса

измерений, и отличает мягкие измерения от мягких вычислений, результаты которых не имеют подобной составляющей.

Рассмотрим основные этапы абстрактной информационно-измерительной технологии формирования решения любой практически реализуемой задачи и определим основные причины и источники формирования неопределенности решения. Выделим следующие этапы:

- подготовки и планирования измерений, включающих подэтапы формирования модели объекта измерения (ОИ), определения измеряемой величины, формирование основных уравнений измерений, а также выбор метода и средства измерений, предварительный выбор алгоритма обработки данных и априорное оценивание погрешности измерений, установление параметров эксперимента, к числу которых относятся объем наблюдений, временные и пространственные точки проведения измерений;

- реализации измерительного эксперимента на основе взаимодействия средств измерений с ОИ и преобразования и сравнения измеряемого и образцового сигналов;

- обработки экспериментальных данных, вычисления метрологически обоснованного результата измерения и его интерпретации согласно цели поставленной измерительной задачи.

На основе этой схемы измерительного процесса можно представить его в целом и отдельные его этапы в формализованном виде уравнениями, записанными в функциональной форме.

Если понимать под обозначениями:

Ψ_j ($j=1, J$) – обобщенные функциональные преобразования, определяющие стратегию реализации этапов;

$\{x_i\}$ ($i=1, I$) – наборы данных, используемых или получаемых на каждом из этапов;

- условия измерений, включающие совокупности априорной информации A , метрологических требований M и ограничений O , то уравнения для каждого этапа и измерительного процесса в целом можно записать в следующем виде:

1. Этап подготовки и планирования измерений Ψ_l , включающий:

- формирование на основе априорной информации моделей ОИ и ВФ $h_l^{(a)}$:

$$h_l^{(a)} = \Psi_l[(x_{l-1} | y_{l-1}) | y_l] \quad (1)$$

$$y_l^{(a)} = (y_{l-1} = \{A_{l-1}, M_{l-1}, O_{l-1}\}) | y_l$$

- определение стратегии измерительного процесса в соответствии с поставленными метрологическими и иными требованиями и ограничениями ($\Psi_{2,l}$):

$$\Psi_{2,l} : \varphi_{j,l} : h_l = \arg \min C[\varphi_{j,l}(\bar{x}_{l-1} | h_l^a; \{b_j\}_l; y_l)], \quad (9)$$

где $\{b_j\}_l \subset B_{j,l}$ – совокупность параметров алгоритма $\varphi_{j,l}$;

- оптимизация параметров измерительного эксперимента ($\Psi_{3,l}$):

$$\Psi_{3,l} : \{b_j\}_l = \arg \min C[\varphi_{j,l}^{(b)}(\bar{x}_{l-1} | \{b_j\}_l; h_l^a; y_l)], \quad (10)$$

где C – критерий оптимизации.

В целом первый этап измерительного процесса можно записать в виде:

$$\Psi_{3,l} : h_l^{(a)} | \{MX\}_l^{(a)} = \Psi_{3,l}(\Psi_{2,l}(\Psi_{1,l}(\bar{x}_{l-1} y_l^a))),$$

где $\{MX\}_l^{(a)}$ – комплекс метрологических характеристик, содержащий показатели качества (точности, надежности, достоверности) полученного решения.

1. Первым этапом решения любой информационной задачи о реальном объекте является этап его предварительного моделирования.

2. Этап реализации измерений $\Psi_l^{(u)}$:

$$\Psi_l^{(u)} = \Psi_{4,l} : \bar{x}_l = \Psi_{4,l}(\{\bar{x}_l\} | h_l^{(a)}; y_l^{(u)}) \quad (11)$$

На этапе реализации измерений вносятся погрешности получения первичной измерительной информации, а также ее обработки по соответствующим алгоритмам, как, например, при косвенных измерениях. В зависимости от типа источника имеется возможность получения как числовых (данные технических измерений, статучетности и другие), так и лингвистических данных и знаний (текстовых описаний, экспертных знаний). В зависимости от типа информации следует использовать соответствующий аппарат описания неопределенностей

либо с использованием функций распределения или ее числовых характеристик, либо функций принадлежности. Это относится как к этапу получения первичных наблюдений или высказываний, так и к этапу обработки по алгоритмам (трансформированные погрешности-неопределенности).

3. Этап обработки и интерпретации результатов измерений $\Psi_l^{(u)}$, состоящий в зависимости от цели задачи из подэтапов.

- - оценивания параметра, аналитического выражения функции или системы функций ($\Psi_{5,l}$):

$$\Psi_{5,l} : h_l^{(o)} | \{MX\}_l^{(o)} = \arg \min C[\varphi_{j,l}^o(\bar{x}_l | h_l^a | y_l^{(o)})] \quad (12)$$

- измерительного контроля ($\Psi_{6,l}$):

$$\Psi_{6,l} : h_l^{(k)} | \{MX\}_l^{(k)} = \arg \min C[\varphi_{j,l}^{(k)}(\bar{x}_l | h_l^a; h_l^{(o)}; y_l^{(k)})] \quad (13)$$

Получения интерпретирующих выводов и решений относительно полученных результатов измерений, являющихся промежуточным результатом или конечной целью задачи; выработка рекомендаций по управлению ОИ или самим процессом измерений на основе полученной экспериментальной информации ($\Psi_{7,l}$):

$$\Psi_{7,l} : h_l^{(y)} | \{MX\}_l^{(y)} = \arg \min C[\varphi_{j,l}^{(y)}(\bar{x}_l | h_l^a; h_l^{(o)}; h_l^{(k)} y_l^{(k)})] \quad (14)$$

В целом, измерительный процесс, включающий все указанные выше этапы можно представить уравнением в следующем виде:

$$h_l | \{MX\}_l = \Psi_l^{(un)}(\bar{x}_l | (\Psi_l^{(u)}(\{\bar{x}_l\} | \Psi_l^{(a)}(\bar{x}_{l-1}; y_l^{(a)}); y_l^{(u)}); y_l^{(un)})) \quad (15)$$

Такая запись отражает формальную основу измерительного процесса как процесса получения метрологически обоснованных знаний на всех этапах измерений и, как будет показано ниже, может служить формальным представлением процесса интеллектуализации измерений.

На этом этапе достигается цель задачи путем применения аппарата вывода окончательного решения в виде (12) – (14). Для первой и второй ситуации предполагается однозначное толкование полученных результатов, не дающее основания для формирования дополни-

тельных неопределенностей. В третьей ситуации при интерпретации возникает ряд дополнительных источников неопределенностей, связанных с как правило субъективной реализацией последнего этапа измерений. В теории и практике принятия решений эти субъективные факторы принимают вид «лиц, принимающих решения». Метрологическое обоснование результатов этого этапа обязательно, иначе погрешности этого этапа могут нивелировать ценность и достоверность всей полученной на предыдущих этапах информации. Эти дополнительные неопределенности в соединении с трансформированными через этот этап неопределенностями первых двух этапов и определяют общую неопределенность результата решения основной задачи.

Очевидно, что для реализации такого рода требований по оптимизации проведения и метрологическому обоснованию всех этапов измерительного процесса необходимы новые методологические и технологические разработки.

1.2.2. Основная концепция байесовской интеллектуализации измерений

Однако, необходимо отметить, что к настоящему времени многие разработки новых направлений в теории измерений еще не завершены и не нашли своего полного практического воплощения.

В связи с этим целесообразно привести результаты, полученные при реализации методологии регуляризирующего байесовского подхода (РБП) и информационных технологий байесовских интеллектуальных измерений (БИИ) в измерительных задачах мониторинга сложных техногенных и природных объектов и принятия управляющих решений. Методологические аспекты РБП и БИИ и реализующие их информационные системы рассмотрены в ряде работ, например [1–17].

Байесовские интеллектуальные измерения, реализующие регуляризирующий байесовский подход, – это измерения, основанные на получении знаний об объекте измерения и использовании их в процессе измерения в совокупности с имеющейся априорной информацией, представленной в виде архивных экспериментальных данных теоретических знаний, аналитических зависимостей, табличной и графической информации, а также рекомендаций и неформализованного опыта специалистов и экспертов. Основное уравнение интеллектуальных измерений (ИИ) записывается в виде (для момента времени $t=t_i$):

$$\begin{aligned} \{h_l \mid MX_l\} &= \{\operatorname{argmin}_C [\varphi_l(x_l \mid y_l)]\}; \\ y_l &= \{A_l; O_l; M_l\}; \\ k &= 1, K; j = J_l; i = 1, I; l = 1, L; \\ h_l &\in H_l, \varphi_l \in \Phi_l, x_l \in X_l, t_l \in T_l, y_l \in Y_l, \end{aligned} \quad (16)$$

где $\{h_l\}$ – список результатов БИИ, апостериорная достоверность каждого из которых определяется значением вероятности P_k ;

H_l – пространство решений – результатов ИИ;

C – оптимизирующее решающее правило (в байесовских интеллектуальных измерениях C – байесовское) выбора решения h_l по алгоритму φ_l из множества алгоритмов Φ_l при наборе экспериментальных данных x_l , метрологических требований M_l , априорной информации A_l и ограничений O_l из соответствующих множеств X_l, M_l, A_l, O_l .

Результатами таких измерений могут быть значения, функциональные зависимости, решения о состоянии объекта и необходимости природоохранных мероприятий, интерпретация экологических ситуаций, обеспеченные полным метрологическим сопровождением в виде комплексов метрологических характеристик, включающих показатели точности, надежности и достоверности.

1.2.3. Новые типы измерений в условиях неопределенности: нечеткие, мягкие, когнитивные измерения

Мягкие байесовские измерения

Одним из новых направлений теории измерений стали *мягкие измерения*, информационных технологии которых позволяют не только* использовать различные типы шкал (метрических и неметрических) для отображения входной (экспериментальной и теоретической) информации (числовой и лингвистической) и формы представления результатов (от полностью определенных до размытых, нечетких, в вероятностной и возможностной интерпретации их достоверности), но и менять саму логику вывода решения в процессе его получения в зависимости от цели задачи и условий ее решения.

В этом типе измерений, в зависимости от типа и условий задачи, может выбираться оптимальная логика вывода решения.

Настраиваемая логика определяется в виде:

$$h_{k,t}(\{l_t \mid MX_{l,t}\}) = \operatorname{argextr}_{C_{t,t}^{(B)}}[\phi_{j,t}^{(l)}(MX_{l,t-1}(h_{k,t-1}; x_{i,t-1}; y_{i,t-1}; l_{t-1}))], \quad (17)$$

где $C_{i,l}^{(B)}$ – критерий выбора логики;

$\phi_{j,t}^{(l)} \in \Phi_{j,T}^{(L)}$ – алгоритм обработки метрологических характеристик, с целью выбора оптимальной логики;

l_{t-1} – логика, при которой было получено решение $h_{k,t-1}$.

Для использования этого механизма нужно перейти к реализации БИИ на спектре настраиваемых логик. При этом обобщенное уравнение таких измерений будет иметь следующий вид:

$$\{h_{k,t}^{(l_t)} | MX_{k,t_1}^{(l_t)}\} = rang \{arg extr C_t[\phi_{j,t}^{(l_t)}((^{(L_T)} * x_{i,t}) * h_{k,t-1}^{(l_{t-1})} | MX_{k_t}^{(l_{t-1})}) | y_{i,t}]\} \quad (18)$$

где $h_{k,t} \in H_{K,T}$ – решение из динамического множества решений на ШДО;

$\phi_{j,t} \in \Phi_{j,T}$ – алгоритм обработки из множества алгоритмов обработки;

$x_{i,t} \in X_{I,T}$ – информационный поток из динамического множества информационных потоков;

$y_{i,t} \in Y_{I,T}$ – набор условий измерений из динамического множества наборов условий измерений;

l_{t-1} – логика, при которой было получено решение $h_{k,t-1}$;

$l_t \in L_T$ – логика из динамического множества логик вывода (в частности, параметрические семейства логик);

$t=1 \dots T$, где T – период времени измерения.

Очевидно, уравнение (11) является обобщением уравнения БИИ. Поскольку настраиваемые логики являются основным элементом теории нечетких систем входящих в направление, получившее название мягкие вычисления, очевидно правомочно назвать данный новый тип измерений *мягкими измерениями МИ (soft measurements SM)*, (рис. 3 иллюстрирует структуру интеллектуальных измерений и место мягких измерений в рамках этого направления) [4], причем если критерий принятия решений останется байесовским, то этот тип уместно назвать *байесовскими мягкими измерениями БМИ (Bayesian soft measurements BSM)*.

В условиях значительной неопределенности, выражающейся в неточности, неполноте, нечеткости знаний и данных о свойствах объекта, среды его функционирования, связей между ними, целей и критериев задач, несогласованностью форм представления и типов информационных потоков, их многочисленностью, малыми объемами статистической информации и сложностью получения их в эксперименте предложен новый

тип реализации измерительного процесса – процессы интеллектуальных и мягких измерений. В этом случае столбцы матрицы представляют собой неслучайные наборы данных. А элементы строк-реперы шкалы имеют характер возможных гипотез о значении измеряемой величины. При этом подходе частотная вероятность заменяется субъективной, «фидуциальной» вероятностью или возможностью реализации гипотез-реперов. При формировании измерительного решения возможными могут оказаться несколько реперов, которые и составят ряд альтернатив. Поскольку каждой гипотезе-реперу соответствует распределение вероятностей или функция принадлежности для каждого набора данных x_i , то матрица преобразуется в куб, в котором третье измерение куба соответствует метрическому пространству возможностей (характеристик уверенности, субъективных или частотных вероятностей) гипотез.

В качестве элементов столбцов могут использоваться лингвистические переменные. В качестве шкал – слабые шкалы: шкалы наименований и порядка, не имеющие вычислительные возможности, но имеющие сильную семантическую наполненность, делающую возможной интерпретацию решений согласно целям измерительной задачи.

В мягких измерениях могут быть реализованы параметрические логики (логика Заде, Лукасевича и другие).

Отличительные свойства МИ можно кратко сформулировать в следующем виде:

1. Измерение реализуется как процесс принятия решений о значении измеряемой величины.
2. В поле решений вводится метрическое пространство возможностей или субъективных вероятностей, значения которых сопровождают измерительный результат.
3. Результат измерений может быть представлен в лингвистической форме.
4. При реализации МИ используются слабые в вычислительном отношении, богатые семантически шкалы (номинальные и порядковые).
5. Результаты МИ представляют собой совокупность альтернатив с метрологическим обоснованием и могут трактоваться как «нечеткие» измерения.
6. Результаты МИ сопровождаются специальными комплексами метрологических характеристик точности, надежности, достоверности, риска, информативности и других.

7. Результатам МИ свойственны активность, интерпретируемость, мотивирующие реализацию мероприятий.

8. МИ реализуются на специальных шкалах (например, шкалах с динамическими ограничениями ШДО [1]), реперами которых представляют собой гипотезы о возможных значениях (градациях) измеряемого свойства.

9. Критерии, логика и правила вывода определяются, исходя из типа измерительной задачи и условий измерений.

10. Шкалы и модели МИ – динамические объекты и могут реформироваться в процессе измерений.

11. МИ применяются, когда нет повторяемости условий реализации измерительного эксперимента, имеются лишь отдельные факты, малые выборки экспериментальных данных.

Можно показать, что уравнения, описывающие концепции компонентов, составляющих направление «мягкие вычисления» (нечеткие системы, нейронные сети, эволюционные вычисления, включающие эволюционные стратегии, эволюционные программы, генетические алгоритмы и генетическое программирование являются частными случаями уравнения МИ. При условии их реализации на принципах единства измерений с полным метрологическим обоснованием получаемых решений очевидно могут быть получены новые виды мягких измерений, таких как нечеткие, нейроэволюционные (генетические) измерения, при условии их реализации на принципах единства измерений с полным метрологическим обоснованием получаемых решений. Так, например, если отдельные составляющие уравнения МИ отсутствуют, в частности, комплексы метрологических характеристик и др., то приходим к уравнениям:

- нечеткие вычисления:

$$\{h_{k,t}\} = \phi_{j,t}^{(l_t)}(x_{i,t}^{(l_t)}) \quad (19)$$

- вычисления в нейронных сетях:

$$h_{k,t} = \phi_{j,t}^{(l_t)}[(x_{i,t}^{(l_t)} * h_{k,t-1}) | y_{i,t}] \quad (20)$$

- эволюционные вычисления:

$$h_{k,t} = \arg \text{extr} C[\phi_{j,t}(x_{i,t})] \quad (21)$$

где $h_{k,t} \in H_{K,T}$ – решение из динамического множества решений на ШДО;
 $\phi_{j,t} \in \Phi_{J,T}$ – алгоритм обработки из множества алгоритмов обработки;
 $x_{i,t} \in X_{I,T}$ – информационный поток из динамического множества информационных потоков;
 $y_{i,t} \in Y_{I,T}$ – набор условий измерений из динамического множества наборов условий измерений;
 l_{t-1} – логика, при которой было получено решение $h_{k,t-1}$;
 $l_t \in L_T$ – логика из динамического множества логик вывода (в частности, параметрические семейства логик);
 $t=1\dots T$, где T – период времени измерения.

Направление мягких измерений может рассматриваться как ветвь направления мягких вычислений, в которых информационные технологии компонентов мягких вычислений реализованы на принципах единства измерений с полным охватом их цепью метрологической поддержки решений. Тогда очевидно, что возможно создание отдельных ветвей мягких измерений в виде:

- нечетких измерительных систем:

$$\{h_{k,t}^{(l_t)} \mid \{MX\}_{k,t}^{(l_t)}\} = \phi_{j,t}^{(l_t)} (* x_{i,t}) \mid y_{i,t} \quad (22)$$

- нейронных измерительных сетей (нейроизмерения):

$$\{h_{k,t}^{(l_t)} \mid \{MX\}_{k,t}^{(l_t)}\} = \phi_{j,t}^{(l_t)} (* x_{i,t}) * (h_{k,t-1}^{(l_t)} \mid \{MX\}_{k,t-1}^{(l_t)}) \mid y_{i,t} \quad (23)$$

- эволюционных измерений:

$$\{h_{k,t}^{(l_t)} \mid \{MX\}_{k,t}^{(l_t)}\} = rang \{arg extr C_t^{(l_t)} [\phi_{j,t}^{(l_t)} (* x_{i,t}) \mid y_{i,t}]\}, \quad (24)$$

где $h_{k,t} \in H_{K,T}$ – решение из динамического множества решений на ШДО;
 $\phi_{j,t} \in \Phi_{J,T}$ – алгоритм обработки из множества алгоритмов обработки;
 $x_{i,t} \in X_{I,T}$ – информационный поток из динамического множества информационных потоков;

$y_{i,t} \in Y_{I,T}$ – набор условий измерений из динамического множества наборов условий измерений.

В работах по мягким вычислениям предлагается классификация структур, реализующих мягкие вычисления, которая иллюстрирует различные сочетания компонентов этого направления. При этом получающиеся в результате комбинации компонентов системы называются гибридными. Очевидным недостатком таких систем являются невозможность интеграции числовой информации в процессе вычислений, отсутствие метрологического обоснования и возможности планирования стратегии реализации процессов получения решений с целью обеспечения их требуемого качества, и в ряде вариантов (нечеткие системы) отсутствие возможности саморазвития на основе самообучения. Структуры, реализующие уравнения (22)–(24), очевидно, будут свободны от указанных недостатков.

Композиция нескольких видов таких измерений будет давать гибридные типы мягких измерений. Построенные на такой методологической основе информационные системы можно назвать *нечеткими измерительными системами, нейронными измерительными сетями, эволюционными измерительными системами и гибридными системами мягких измерений* (при их композиции) [4].

Объединение систем мягких вычислений с системами мягких измерений может производиться на нескольких уровнях интеграции, которые будут обеспечивать различную глубину взаимного проникновения методологий [1–5].

Концептуальная модель МИ может быть формализована на основе следующих рассуждений.

Как известно, в процессе измерений система объекта измерений G_x преобразуется в систему модели ОИ G_{mx} в одностороннем порядке при гомоморфном преобразовании Φ вида []:

$$G_x \rightarrow \{G_{mxi} | P(G_{mxi})\} : \Phi(Q_{mxi}) = x_i \quad (25)$$

где Q_{mx} – измеряемое свойство ОИ;

$\{G_{mxi} | P(G_{mxi})\}$ – мягкая модель объекта измерения в виде совокупности альтернативных моделей с советующими им вероятностями $P(G_{mxi})$;

$i=1, I$;

x_i – результат измерения по i -той модели.

В процессе МИ измерительное преобразование производится на основе правил вывода и принятия решений Фми. Его применение приводит к формированию альтернативных результатов измерений свойств этих моделей, что может быть концептуально представлено в виде:

$$G_x \rightarrow G_{mx} \rightarrow \Phi_{mi}(Q_{mx}) = \{x_i\}, I=1, \dots, I, \quad (26)$$

где I – число альтернативных результатов измерений.

Принятие решений в условиях риска и неопределенностей связано с данными, задаваемыми обычно либо в матричной форме, либо в форме совокупности измеряемых свойств $\{Q_{imx}\}$ ОИ, описываемых распределениями вероятностей или функций принадлежности.

Рассмотрим матричное представление измерительного процесса при реализации функции измерительного преобразования (сопоставления) на основе аппарата принятия решений в условиях ограниченной неопределенности, когда известен риск альтернативных решений.

Строки матрицы V_{ik} соответствуют возможным альтернативным значениям результатов измерений h_k , столбцы – наборам исходных экспериментальных данных $\{x_i\}$. Если определено расстояние между элементами h_k , строка матрицы представляет собой измерительную шкалу S с реперами h_k . Элемент матрицы v_{ik} может быть интерпретирован как плата за решение о выборе репера h_k -го при появлении набора данных x_i . Матрица V_{ik} при этом называется платежной. Если же элементы матрицы представляют собой выигрыш, полезность, доходы, прибыль, то матрица носит название полезности.

В качестве результатов могут выступать элементы матрицы риска, которые связаны с платой или выигрышем соотношением:

$$\begin{aligned} V_{ki} - V_i, & \text{ если } V - \text{плата, затраты;} \\ R_{ki} = V_i - V_{ki}, & \text{ если } V - \text{выигрыш, доходы.} \end{aligned}$$

Решении о выборе некоторого h_k при условии реализации x_i принимается на основе определенных критериев или правил выбора.

В условиях рисков при ограниченной неопределенности в качестве таких критериев и правил могут быть выбраны: правило Лапласа, правило Байеса в классической форме (при известных априори компонентах формулы Байеса), критерий среднего и стандартного отклонения, критерий Вальда (минимаксный, максимаксный), критерий

Сэвиджа (величина платы лица, принимающего решения), критерий Гурвица и другие.

При таком подходе классические измерения могут быть представлены как частный случай, при котором матрица V_{ik} принимает вид единичной (для прямых измерений) или диагональной (для косвенных или функциональных измерений).

Важно, что при этих условиях число и вид элементов матрицы определены или могут быть точно определены в измерительном эксперименте, то есть если и имеется неопределенность, то она может быть снята путем получения дополнительной информации в процессе измерений. Этот подход может быть реализован в виде адаптивных, алгоритмических измерений.

Если элементы столбцов x_i представляют собой выборки случайных процессов или величин, то подход может быть реализован в виде статистических измерений.

В этом случае элементы столбцов матрицы имеют стохастический характер и могут быть представлены выборками статистических значений экспериментальных данных или функциями от них, например функциями распределений вероятностей. Элементы строк (реперы шкалы) имеют неслучайный характер.

1.2.4. Когнитивные измерения на основе байесовских интеллектуальных технологий

Расширение понятия «измерение» в методологии БИТ как в методологическом направлении обобщенной теории измерений, так и в информационно-технологическом, привело к созданию новых измерительных средств, способных измерять параметры сложных объектов в условиях значительной неопределенности информации о них.

Дальнейшее развитие этого направления теории измерений связано с вовлечением субъективной составляющей (измерителя) в измерительный процесс в качестве источника информации и в качестве аналитического «решателя» измерительной задачи.

В качестве «измерителей» могут быть лица, принимающие решения, эксперты, пользователи и другие лица.

Позиционирование измерителя в качестве источника измерительной информации позволяет получить дополнительные знания об объекте измерения, среде его окружения и условиях измерений (измерительных средствах, метрологических требованиях, имеющейся априорной информации, постановке измерительной задачи, ограничениях

и допущениях моделей и подобной информации). Как правило измеритель имеет значительный запас подобных знаний и, кроме того, умений настраивать все компоненты измерительного процесса в соответствии с требованиями задачи.

В традиционных измерениях физических величин в условиях определенности в функции измерителя входила деятельность по аналитической оценке и интерпретации результатов измерений. В целях повышения эффективности и расширения функциональности измерительных процессов эти знания и умения необходимо использовать на новой информационно-технологической основе, обеспечивающей их интеграцию с результатами измерительных приборов и систем обработки информации.

Алгоритмы и технологии искусственного интеллекта позволяют сделать возможным извлечение знаний субъекта. Этому направлению посвящены многочисленные работы отечественных и зарубежных ученых в сфере искусственного интеллекта и мягких вычислений [4–7].

Активно развивающееся направление когнитивных технологий целиком направлено на решение этой задачи. Однако, в отличие от когнитивной графики, эти технологии находятся еще в начале разработок и не дают возможности на формальной основе использовать их решения в измерительном процессе, разрабатывать технологии и системы.

Основная сложность привлечения в измерительный процесс и формализации знаний и умений измерителя состоит в необходимости их метрологического обоснования и технологической интеграции в измерительный процесс.

Знаниям и умениям измерителя, как субъективным и относительным, всегда присуща значительная степень неопределенности, сложности противоречивости. При использовании таких знаний в измерительном процессе возникает дополнительная энтропийная составляющая. Это приводит к появлению или повышению степени неустойчивости и некорректности постановки задачи, снижает эффективность и качество решений или делает невозможным получение решений в требуемой постановке измерительной задачи.

Субъектов-измерителей может быть несколько. Они могут быть распределены в пространстве и во времени (например, исторические сведения конкретных субъектов) и образовывать информационную сеть. Такая сеть, будучи интегрирована с сетью измерительных систем задачи, может вносить значительную энтропийную составляющую

щую, которая будет способна создать хаотичность в измерительном процессе и обусловить кризисные ситуации, что особенно важно в задачах измерительного управления, когда измерительные решения являются управленческими. В практике наблюдается подобное явление при управлении сложными технологическими или социально-экономическими комплексами субъектами, недостаточно квалифицированными, имеющими противоречивые мнения и знания или собственные интересы, не соответствующие цели задачи. Примерами тому являются экономические или техногенные кризисные ситуации.

С другой стороны, метрологически аттестованные (квалиметрированные) и формализованные в виде моделей и технологий знания и умения субъектов при свертке их с информацией измерительных систем могут обеспечить не только значительное повышение эффективности решений и усиление мощности измерительных систем, но и решение прежде нерешаемых измерительных и управленческих задач. Развиваясь во времени и в пространстве, наращивая число субъектов-измерителей – агентов сети, такая измерительная сеть может стать реальной **сетью измерительных знаний** для решения важнейших и сложнейших задач мониторинга, контроля и управления современной экономики, охраны окружающей среды, науки, образования, культуры.

В данной работе предлагается методологическая основа и технологии реализации измерительного процесса, в технологический контур которого вовлечены один или несколько субъектов-измерителей в качестве источников информации (знаний и умений) и распределенного аналитического решателя (блока принятия решение и аналитического вывода). Таким образом, **измерения, в технологиях которых на формальной основе извлекаются и используются метрологически аттестованные знания и умения субъектов, называются когнитивными измерениями.**

Для реализации когнитивных измерительных процессов в условиях значительной неопределенности может быть использован регуляризирующий Байесовский подход (РБП) и байесовские интеллектуальные технологии (БИТ).

Основными базисными понятиями РБП и БИТ являются динамический компакт решений (ДКР), модель с динамическими ограничениями (МДО) и шкала с динамическими ограничениями (ШДО). Именно на их основе создаются новые измерительные технологии.

Рассмотрим их применение для реализации когнитивных измерений.

Топологическая модель пространства решений когнитивных измерений должна представлять собой динамический компакт в определении [1], представляющий собой ограниченное метрическое функциональное пространство G_{dst} , для которого определено понятие расстояния d с метрикой для момента времени t . Элементами топологического пространства компакта решений являются результаты когнитивных измерений. Элементами топологического пространства **когнитивной измерительной сети** G_{st} являются носители знаний и умений (субъекты-измерители), причем пространства (компакты) решений и измерителей связаны друг с другом изоморфным измерительным преобразованием:

$$G_{st} \rightarrow G_{dst} \quad (27)$$

Связанность компактов определяет их взаиморазвитие, когда добавление элементов компакта G_{st} приводит к расширению компакта G_{dst} .

Компакт решений задачи G_{dot} с учетом измерительных систем определяется сверткой компактов решений субъектов G_{dst} и систем G_{dmt} .

$$G_{dot} = G_{dst} * G_{dmt} \quad (28)$$

Модель с динамическими ограничениями Q_{ist} для информационного потока, получаемого от субъекта как источника информации или как аналитика-решателя, определяющего решение из некоторого списка альтернатив или корректирующего информационно-измерительный процесс, определяется совокупностью свойств субъекта Q_{st} и его связей с окружающей средой R_{set} .

$$Q_{ist} = \{Q_{st}, R_{set}\} \quad (29)$$

Свойства субъекта представляют, как минимум, три группы свойств: свойства субъекта как специалиста в решении данной задачи Q_{sqt} , свойства субъекта как личности Q_{spt} , имеющие значение для решения данной задачи, свойства субъекта в иерархической структуре управления процессом решения Q_{smt} , определяющие важность, значимость и степень влияния его решений на итоговое решение задачи. Модель с динамическими ограничениями, учитывающая эти группы свойств может быть представлена в виде:

$$Q_{ist} = Q_{sqi} * Q_{spt} * Q_{smt} \quad (30)$$

где символ * определяет байесовскую свертку, определенную в [1].

Знания субъекта представляются в виде информационных массивов определенного формата, а умения в виде технологий, формализованных на основе БИТ.

Определив состав и структуру свойств субъекта, можно приступить к формированию измерительных шкал ШДО для измерения данных свойств. Для измерения количественных характеристик можно использовать числовые ШДО. Например, для измерения КПД котла теплогенерирующей организации конструкция ШДО приведена на рис. 10.

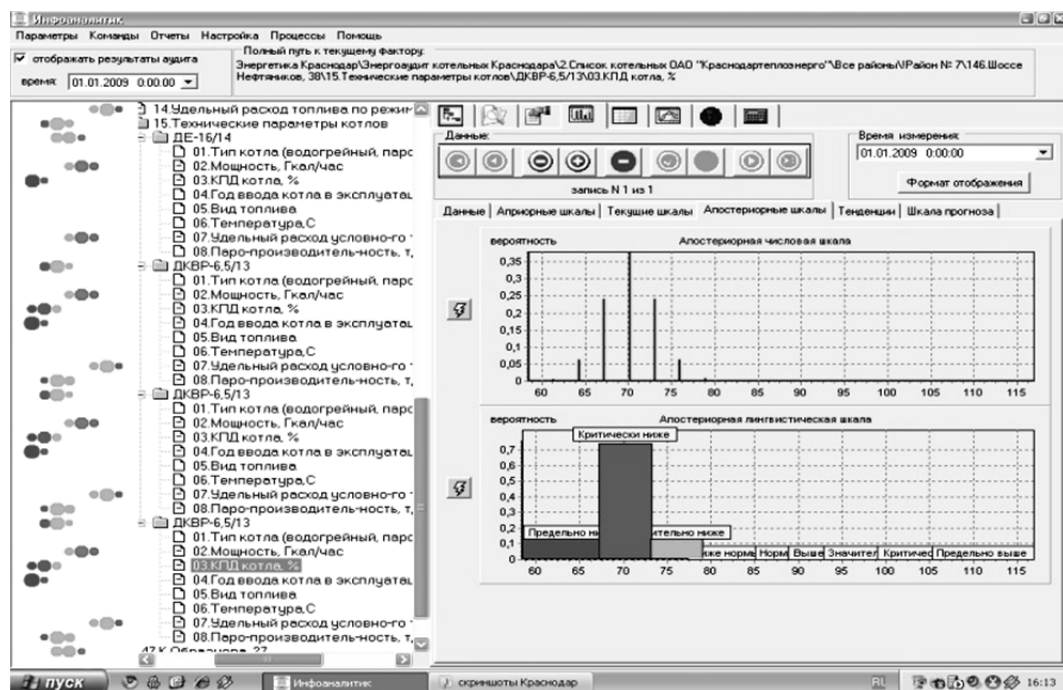


Рис. 10. Пример измерения КПД котла теплогенерирующей организации

Для измерения качественных показателей рекомендуется использовать лингвистические ШДО. Примером может служить шкалы ШДО для измерения значимости решений субъекта и его личной заинтересованности в решении данной задачи (рис. 11, 12).

Извлечение знаний и умений субъекта по методологии БИТ реализуется средствами платформы «Инфоаналитик». «Инфоаналитик» представляет собой программную среду, предназначенную для быстрой раз-

работки приложений, которыми могут быть экспертные системы различного назначения, системы поддержки принятия решений; системы мониторинга сложных объектов, процессов, ситуаций; системы контроля и аудита, системы управления, многокритериальные системы и другие. Основные компоненты регуляризирующего байесовского подхода (ДК, МДО, ШДО) реализованы в «Инфоаналитике» в виде объектов, которые создаются пользователем средствами создающихся комплекса программных модулей, имеющих усиленную поддержку интеллектуальным интерфейсом и когнитивной графикой. Примеры когнитивной графики, используемой в процессе аудирования состояний котельных приведены на **рисунках 11, 12** в виде разноцветных кружков, отражающих степень соответствия показателей объекта их нормативным значениям. В подсистеме когнитивной графики используется 9 цветов, отражающих как опасные, так и благоприятные состояния объектов.

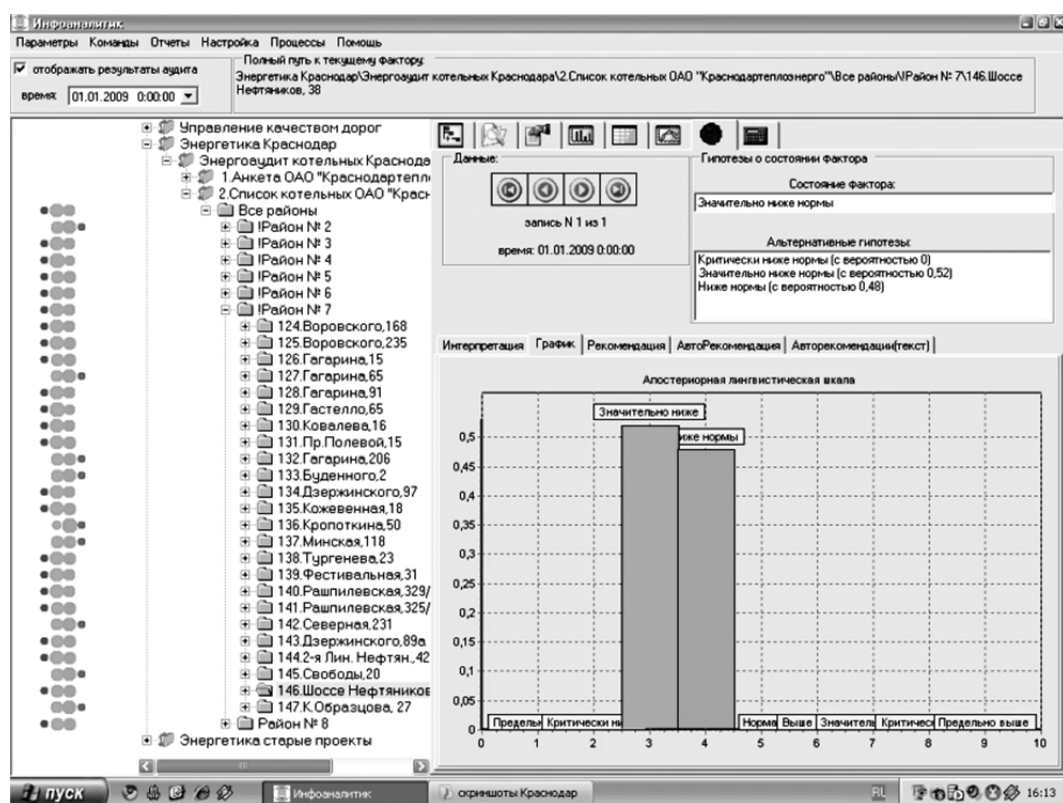


Рис. 11. Пример лингвистической оценки общего состояния котельной

Формирование модели объекта и среды его функционирования может производиться разработчиками или пользователями предварительно или в процессе функционирования системы.

При извлечении знаний об объекте и среде измеритель либо корректирует предложенное дерево факторов, отражающих свойства объекта, существенные для решения задачи, либо создает его заново. Интеллектуальный интерфейс предлагает измерителю для заполнения простые формы для установления связей между свойствами, типа их влияния, их значимости. Пример такой модели приведен на **рисунке 11** для конкретной прикладной задачи.

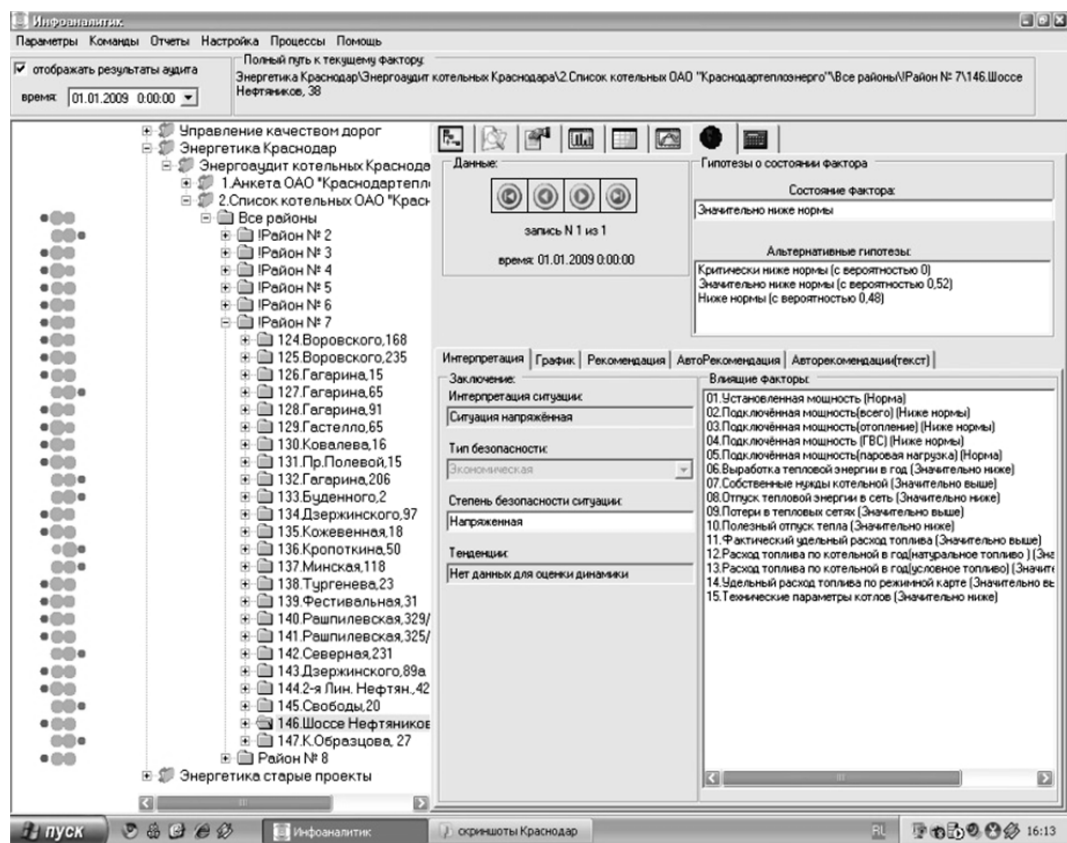


Рис. 12. Интерпретация состояния и влияющих факторов среды функционирования котельной

Затем детализируется модель объекта, при этом знания получаются от субъекта о значениях или оценках значений свойств объекта и среды, как в числовой, так и в лингвистической форме. Этот процесс связан с созданием субъектом различных ШДО, соответствующим свойствам объекта и среды, которые в совокупности формируют гиперкуб динамического компакта решений. Пример ШДО приведен для оценки свойств котельных.

При формировании ШДО субъекту предлагается создать или обновить нормативно-методическую базу, которая хранится в модуле «Элек-

тронная библиотека» используется при создании ШДО. Таким образом извлекаются знания субъекта оценочного и аудлирующего характера.

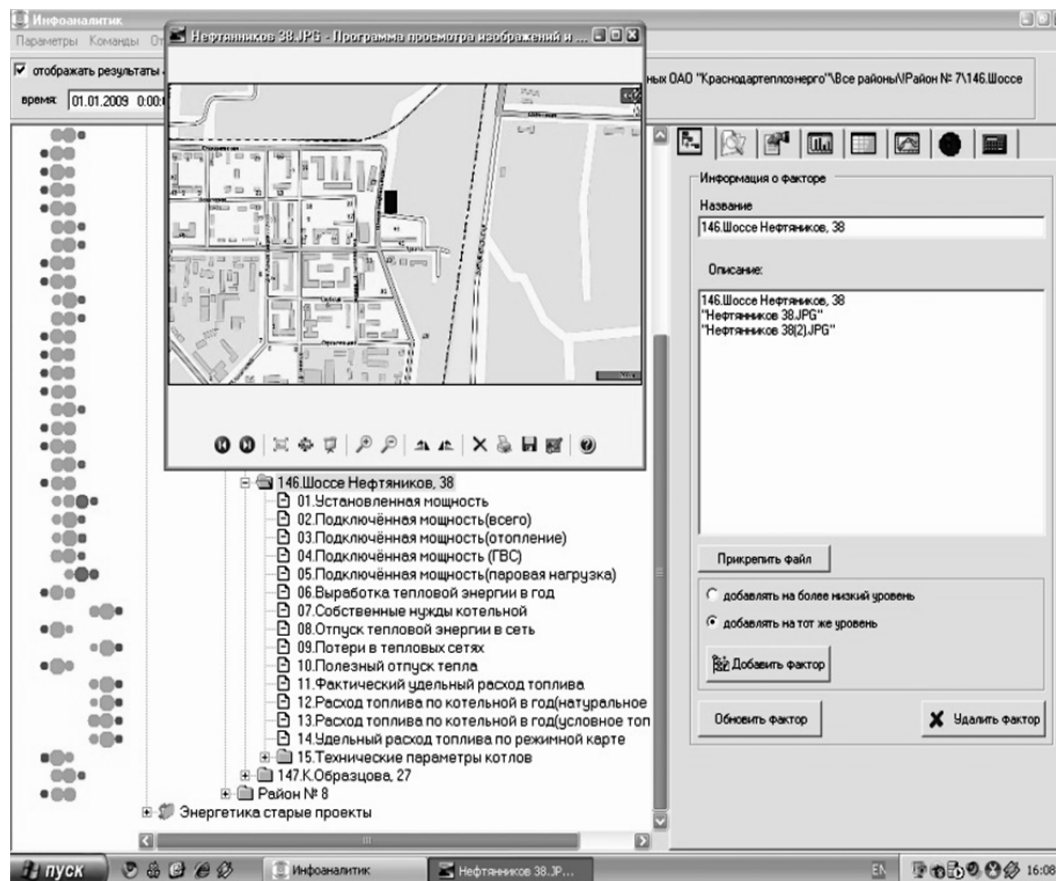


Рис. 13. Модель сети котельных и среды их функционирования

Все эти средства позволяют извлекать знания субъекта.

Умения субъекта отражаются в виде технологий или алгоритмов решения задачи или подзадач. Для извлечения умений субъекта используются специальные вычислительные модули платформы «Инфоаналитик»: «Калькулятор» для вычислений регуляризированных байесовских оценок (РБО), библиотека байесовской математической статистики (БМС), блок генерации рекомендаций.

Все эти средства позволяют извлекать знания субъекта.

Умения субъекта отражаются в виде технологий или алгоритмов решения задачи или подзадач. Для извлечения умений субъекта используются специальные вычислительные модули платформы «Инфоаналитик»: «Калькулятор» для вычислений регуляризированных байесовских оценок (РБО), библиотека байесовской математической статистики (БМС), блок генерации рекомендаций

Модуль калькулятор [3] позволяет субъекту ввести в состав алгоритмической базы «Инфоаналитика» новую функциональную зависимость и применить ее к обработке как четких, так и нечетких переменных, представленных в виде РБО. При этом функциональная зависимость приобретает вид алгоритма и может активно использоваться. Косвенные или аналитические измерения реализуются на этой основе.

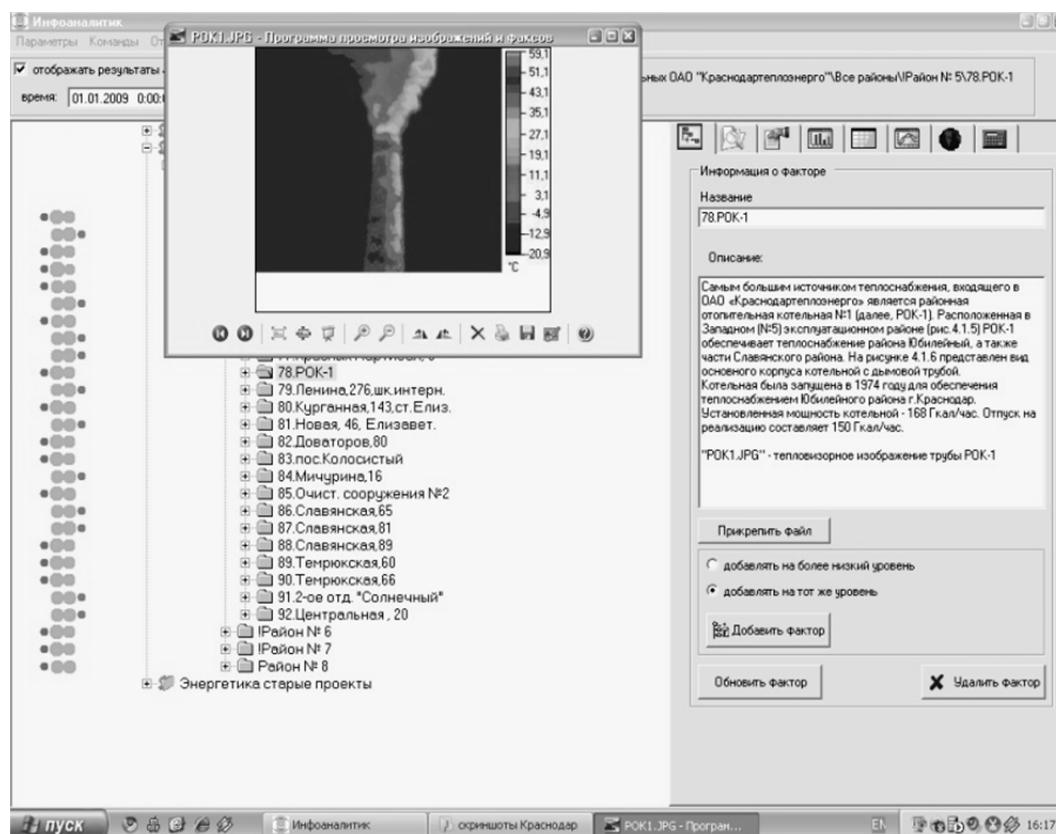


Рис. 14. ШДО для измерения утечки тепла на основе свёртки фактически измеренных данных

Библиотека байесовской математической статистики, подробно описанная в [5], служит для обработки выборок (прежде всего малых выборок) статистических данных. Она содержит как алгоритмы классической математической статистики, так и алгоритмы байесовской математической статистики, работающие в условиях малых выборок и значительной неопределенности. В качестве данных для БМС могут быть выборки, состоящие не из вероятностных, а из «возможностных» данных, для которых вероятность их появления заменяется возможностью или фидуциальной (субъективной) вероятностью. Это да-

ет возможность обрабатывать ряды данных, сформированные субъектом на основе его некоторых предположений.

Кроме того, БМС позволяет субъекту самому сформировать не только выборку данных, но и процесс их обработки. Для этого он может извлечь из библиотеки БМС отдельные алгоритмы и объединить их в технологию, задавая необходимую последовательность их реализации в вычислительном процессе. Средствами БМС реализуются алгоритмы статистических измерений.

Модуль генерации рекомендаций позволяет пользователю задать некоторые коррективные действия по усилению или ослаблению влияния какого-либо фактора, создавая тем самым управляющее воздействие в измерительно-аналитической системе.

Прежде, чем использовать знания или умения субъекта необходимо их метрологически аттестовать, то есть поставить в соответствие каждому информационному элементу (массиву информации или технологии) комплекс метрологических характеристик.

В методологии БИТ комплекс метрологических характеристик $\{MX\}$ представлен показателями точности, надежности и достоверности, а также риска решений и их энтропийных характеристик, что позволяет эффективно управлять измерительным процессом в ходе эксперимента. Решение субъекта может быть определено по оптимизационному уравнению Бит в виде:

$$\{h / \{MX\}\}_{dst} = \arg \min C_t[f_t(\{x_{it}\} / y_{it})] \quad (31)$$

Однако для его практического применения в информационно-измерительном процессе необходимо свернуть его решение с оценкой его свойств, точнее метрологические характеристики его решения $\{MX\}_{dst}$ с метрологическими характеристиками оценок его свойств $\{MX\}_{fst}$

$$\{MX\}_{dst} = \{MX\}_{dst} * \{MX\}_{fst} \quad (32)$$

Аттестованные решения каждого субъекта сворачиваются методом Байесовской свертки для получения общего решения субъектов в виде:

$$\{MX\}_{dsbt} = \{MX\}_{dsbt} * \{MX\}_{dsit} \quad (33)$$

Таким образом можно получить взвешенное коллективное решение специалистов по определенному вопросу организовать работу с экспертами и дискуссии в социальных сетях и на общественных слушаниях. Для получения решения измерительной задач в целом $\{h|\{MX\}\}_{dt}$ полученное коллективное решение должно быть свернуто с измерительной оценкой, полученной средствами измерений и систем обработки информации $\{h|\{MX\}\}_{dsmt}$.

$$\{h / \{MX\}\}_{dt} = \{h / \{MX\}\}_{dsbt} * \{h / \{MX\}\}_{dsmt} \quad (34)$$

Для извлечения знаний должны быть построены специальные ШДО предметной области, примеры которых реализованы в работах автора и его коллег и приведены, например, в [2–4]. Для формализации умений субъекта используется инструментальная база БИТ в виде платформы «Инфоаналитик», реализованная для формализации и получения знаний в различных предметных областях. Для конкретных практических приложений разработаны прикладные системы в виде интеллектуальных рабочих мест специалистов ИРМ [5–7], которые позволяют создавать базы знаний и прикладных методов и технологий решения измерительных задач в условиях значительной неопределенности.

1.3. Интеллектуальные информационные технологии на основе регуляризирующего байесовского подхода

1.3.1. Интегрирующие технологии на основе байесовских интеллектуальных измерений

К настоящему времени созданы многочисленные разноплатформные базы данных, расчетные методики и модели, распределенные по научным, ведомственным и производственным организациям. Системный подход к оценке ситуаций и состояний объектов природной и техногенной сред обуславливает необходимость интеграции модельных представлений различных физических и информационных процессов и явлений. Так, например, при определении степени надежности сложного техногенного комплекса (предприятий по добыче и переработке природных ресурсов, линий электропередач, газонефтепроводов, атомных электростанций, теплоэлектростанций и др.) необходимо интегрировать модели влияния окружающей природной

среды (геодинамические, гидрологические, климатические процессы) и ее техногенной насыщенности (виды и степень надежности эксплуатируемых технических объектов, наличие мест выработки полезных ископаемых, строительство новых сооружений), а также характеристик социально-экономической сферы. Однако, несмотря на имеющиеся перспективные технологии по извлечению и транспортировке информации и типизированность основных прикладных задач оценивания, контроля и управления, для получения практически приемлемого решения конкретной прикладной задачи каждый раз необходима достаточно длительная и трудоемкая работа по созданию информационной технологии решения задачи в целом, обеспечивающей сбор и согласование имеющихся данных и моделей, а также реализацию расчетных и аналитических этапов, которая выполняется в рамках специального проекта группой специалистов-разработчиков. Многие этапы такой работы могли бы быть с успехом автоматизированы на основе применения интегрирующих технологий. Первым шагом в этом направлении явились технологии CLIENT-SERVER, обеспечивающая на основе серверов баз данных быстрый доступ пользователей к распределенным информационным ресурсам. Однако возможности таких технологий в настоящее время весьма ограничиваются недостаточностью интеграции только данных (а не совокупности данных и знаний), невозможностью управления качеством и риском получаемых решений, неавтоматизированностью процесса адаптации и развития технологий в меняющихся информационных условиях, например, при добавлении нового источника информации (нового измерительного прибора, новой базы данных, новых модельных представлений).

В работах научной школы РБП приводятся убедительные примеры эффективности БИИ и БИТ-технологий при решении широкого круга задач экономики, природопользования, распределенного индустриального производства, финансов, охраны окружающей среды (например, доклады А.С. Пиотровского, Р.А. Фокина, О.Г. Иванова, А.В. Ивановой, С.Ж. Филиппова, С.И. Крамарь, С.В. Бровчака и др.).

Системные или ситуационные БИТ-технологии предназначены для выявления типа и интерпретации ситуаций, а также сценариев их развития на основе выявления и свертки многочисленных влияющих на состояние объектов факторов внешней среды. Решение в виде регуляризированной байесовской оценки ситуации получается на основе следующей зависимости:

$$h_{kj}^{(cou)} = h_{kj}^{(ou)} * h_{kj}^{(mou)},$$

где $h_{kj}^{(cou)}$, $h_{kj}^{(ou)}$, $h_{kj}^{(mou)}$ – сопряженные РБО ситуации, параметров и результатов моделирования на основе комплекса моделей $\{ \Pi_2^{(i)} \}$, 7 – число интегрированных по БИТ-технологиям расчетных моделей.

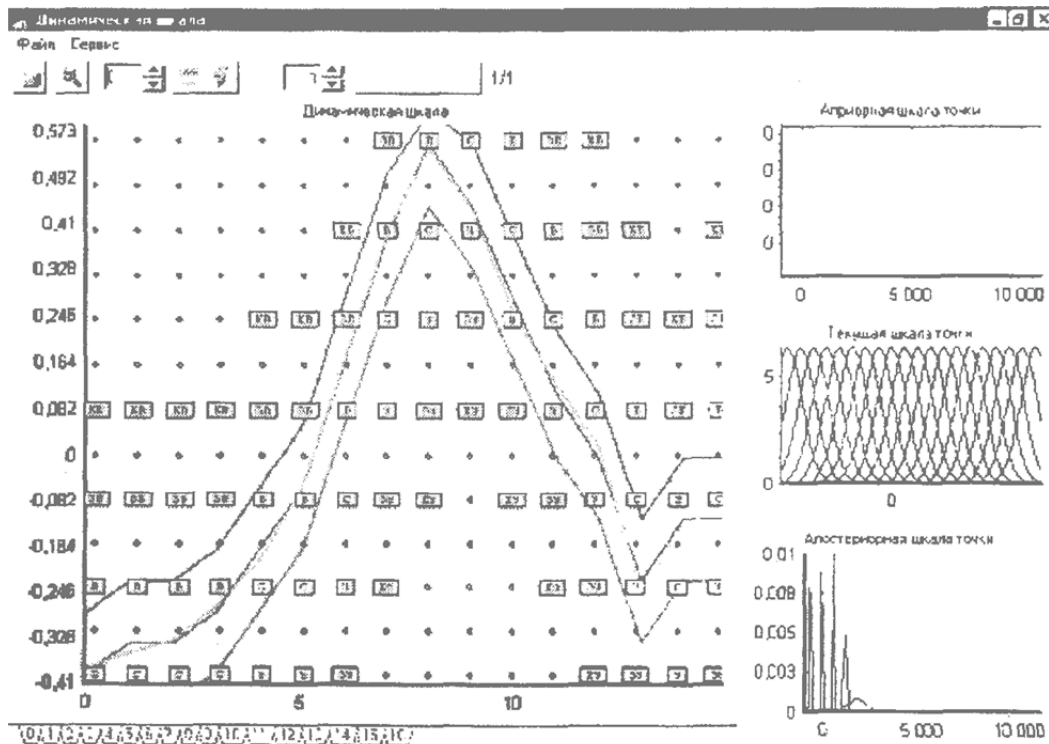


Рис. 15. Нечеткая модель корреляционной функции на основе БИТ

На рисунке 16 иллюстрируется оценка и интерпретация экологического состояния воздушной среды промышленного региона. Определяется степень безопасности состояния, тенденции ее развития, альтернативные оценки ситуации (нечеткое оценивание), значимо влияющие факторы, рекомендации по возможным способам ее корректировки. Можно рассмотреть влияние факторов детальнее, если перейти на предыдущие нижние уровни дерева факторов и тем самым полнее интерпретировать причины возникновения данной ситуации.

Результатом реализации БИТ-технологий на уровне генерации управляющих рекомендаций являются гибкие стратегии управления производством и использованием ресурсов.

Принципиальным свойством БИТ-технологий является параллельное с процессом основных преобразований информации непре-

рывное метрологическое обоснование, что придает им измерительный характер. Для функциональных (модельных) и системных (ситуационных) БИТ-технологий на каждом этапе этих технологий производятся преобразования комплексов метрологических характеристик на основе формализованного подхода, приведенного в [1].

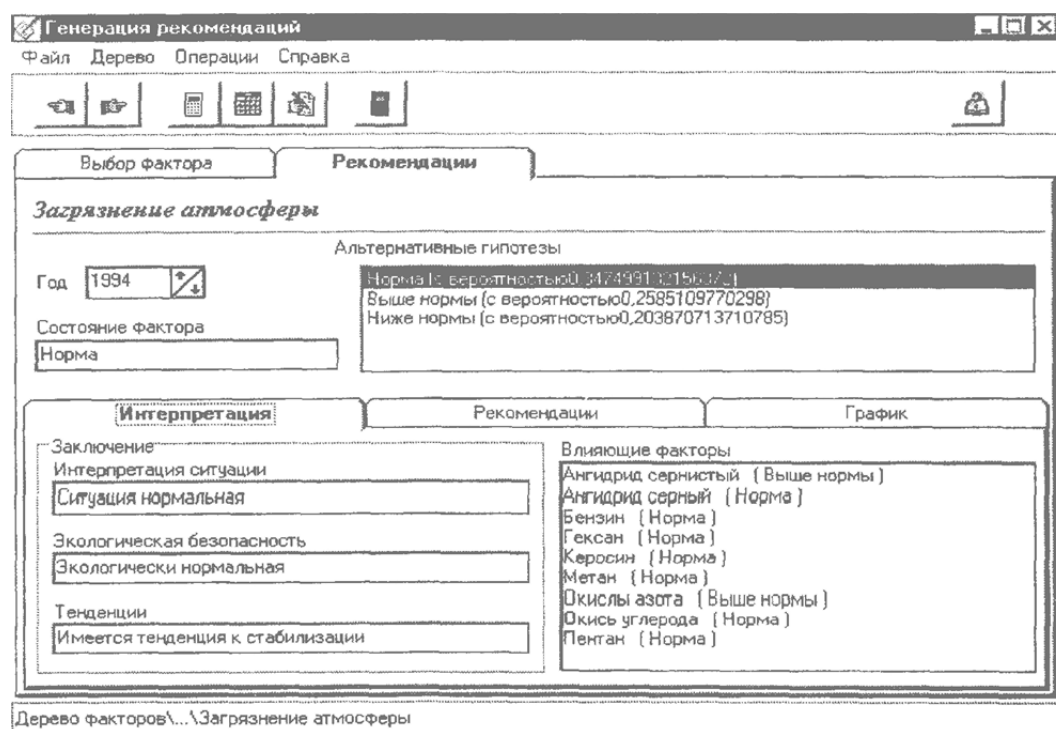


Рис. 16. Интерпретация ситуации на основе БИТ-технологий

Формализованная метрологическая основа БИТ-технологий позволяет определять степень практической приемлемости получаемых оценок, прогнозов, рекомендации, обеспечивает возможность управления качеством решений на основе согласованного решения прямой и обратной задач метрологии БИИ и БИТ. Эти же свойства БИТ-технологий создают принципиальную основу и методическую базу для задач инвестирования и оценки риска новых проектов, реконструкции предприятий, расширения производств, строительства сооружений, развития территорий.

1.3.2. Принципы создания развивающихся информационных технологий на основе регуляризирующего байесовского подхода

Одним из актуальных направлений развития современной информационной индустрии является разработка новых подходов и методоло-

гий построения гибких информационных технологий для управления сложными объектами в условиях меняющихся внешних ситуаций, собственных характеристик объекта и взаимосвязей между ними. Такие технологии не имеют жесткой алгоритмической структуры и могут развиваться автономно или пользователем по мере изменения требований, ограничений или целей решаемых задач. Прикладные задачи такого типа характеризуются значительной информационной неопределенностью, выражающейся в неточности, неполноте, нечеткости знаний и данных о свойствах объекта, среды его функционирования, связей между ними, целей и критериев задач, несогласованностью форм представления и типов информационных потоков, их многочисленностью, малыми объемами статистической информации и сложностью получения их в эксперименте. Это затрудняет и делает неэффективным применение классических подходов к управлению, основанных на принципах использования в процессе управления заранее созданных моделей объектов и алгоритмов их функционирования.

В настоящее время разрабатывается достаточно много подходов к созданию методологической основы и технологий обработки информации в условиях неопределенности на базе методов искусственного интеллекта.

Как показано в ряде работ, [1–3] такие подходы реализуются в рамках концепций построения самоорганизующихся моделей и методов. Однако, в целом требования прикладных задач к управлению синтезом, качеством и эффективностью решений не обеспечиваются в немногочисленных *практических реализациях* предлагаемых методических разработок. Объясняется это тем, что разработчики методов искусственного интеллекта, уделяя значительное внимание формам представления информации (в основном нечисловой) и семантической интерпретации результатов, практически не прорабатывают вопрос синтеза информационных технологий *как единого целостного процесса, реализуемого в рамках требуемых показателей качества и эффективности получения решений*. Кроме того, концепции моделей объектов построены на принципах жесткого априорного конфигурирования, даже если в них и заложены динамические показатели. Поэтому задачу построения нового класса информационных технологий, очевидно, необходимо начать с концептуального рассмотрения методологических аспектов создания моделей сложных объектов, которые являются содержательной частью и составляют самую суть алгоритмических построений.

В условиях значительной неопределенности модель сложного объекта и среды его окружения должна изменяться в зависимости от получаемой информации и меняющихся требований, ограничений, целевых функций и критериев выполняемой задачи. В концептуальной форме это можно определить как изменение степени «погружения» системы модели $G^{(M)}(t)$ в систему объекта $G^{(O)}(t)$ и формально представить в виде [1] гомоморфного отображения:

$$G^{(O)}(t) \rightarrow G^{(M)}(t), \quad (35)$$

где $G^{(O)}(t)$ – система динамического объекта $G^{(O)}(t) = Q^{(O)}(t) * R^{(O)}(t)$ со свойствами $Q^{(O)}(t)$, отношениями $R^{(O)}(t)$, меняющимися в зависимости от времени t ;

$G^{(M)}(t)$ – система динамического объекта $G^{(M)}(t) = Q^{(M)}(t) * R^{(M)}(t) * L^{(M)}(t)$ со свойствами $Q^{(M)}(t)$ и отношениями $R^{(M)}(t)$ и ограничениями, допущениями, требованиями $L^{(M)}(t)$ постановки задачи, также меняющимися во времени.

Для природных и техногенных объектов, активно взаимодействующих с природной средой, это погружение носит бесконечный характер в силу принципиальной невозможности получить исчерпывающую информацию о природных объектах, что может быть выражено для независимых составляющих в следующем виде:

$$H^{(M)}(t) = H^{(O)}(t) + \Delta H^{(M)}(t), \quad (36)$$

$$|\Delta H^{(M)}(t)| > 0,$$

где $H^{(O)}(t)$ – энтропия системы объекта;

$H^{(M)}(t)$ – энтропия системы модели;

$\Delta H^{(M)}(t)$ – остаточная энтропия системы, неснимаемая в результате решения задачи в данной постановке при использовании модели и соответствующей ей информационной технологии.

Энтропия системы модели может быть представлена как сумма энтропий систем свойств $H^{(Q)}(t)$, отношений между ними $H^{(R)}(t)$, полноты самой системы свойств $H^{(QQ)}(t)$ и полноты системы отношений $H^{(RR)}(t)$, а также энтропий системы ограничений $H^{(L)}(t)$ и ее полноты $H^{(LL)}(t)$. Эти составляющие зависят от меняющейся информационной

ситуации (наличие экспериментальных данных, их точности, объёма и др.) с присущей ей степенью неопределённости $H^{(I)}(t)$ и неопределённостью в постановке (конкретной модели, целей, критериев, ограничений) прикладной задачи $H^{(Z)}(t)$. С учётом этого уравнение (36) можно представить в виде:

$$H^{(M)}(t) = H^{(O)}(t) + \Delta H^{(I)}(t) + \Delta H^{(Z)}(t), \quad (37)$$

где $\Delta H^{(Z)}(t)$ – энтропия в постановке задачи, которая не снимается в процессе накопления данных, если система моделей $G^{(M)}(t)$ не меняется. Таким образом, для обеспечения адекватности модели объекту и снижения неопределённости необходимо постоянное развитие модели и коррекция самой постановки задачи в режиме мониторинга условий ее решения.

Если учесть, что эти составляющие могут быть взаимозависимыми, то, очевидно, в уравнениях (36), (37) появятся составляющие, отражающие в энтропийном смысле их взаимодействие. Тогда уравнения могут быть записаны в виде уравнений байесовской свертки, которые для двух составляющих могут быть записаны в виде:

$$\Delta H^{(M)}(t) = \Delta H^{(Z)}(t) * \Delta H^{(I)}(t) = \Delta H^{(Z)}(t) | \Delta H^{(I)}(t) * \Delta H^{(I)}(t) \quad (38)$$

Для K взаимосвязанных источников неопределённости байесовская свертка вида (38) реализуется по формуле:

$$H^{(i)}(t) | H^{(j)}(t) = \frac{H^{(i)}(t) * H^{(j)}(t) | H^{(i)}(t)}{\sum_{k=1}^K H^{(i)}(t) * H^{(k)}(t) | H^{(i)}(t)} \quad (39)$$

Детализация данных формул для построения иерархических структур энтропийных представлений модели согласно уравнений (36)–(39) дана в работе [1].

В соответствии с вышесказанным модель сложного объекта во взаимодействии со средой ее функционирования в энтропийной интерпретации может быть записана в виде:

$$H^{(MOE)}(t) = H^{(MO)}(t) * H^{(ME)}(t) = H^{(MO)}(t) | H^{(ME)}(t) * H^{(ME)}(t), \quad (40)$$

$$\text{где } H^{(ME)}(t) = Q^{(ME)}(t) * R^{(ME)}(t) * L^{(ME)}(t) \quad (41)$$

энтропия при моделировании свойств, отношений и ограничений среды функционирования объекта.

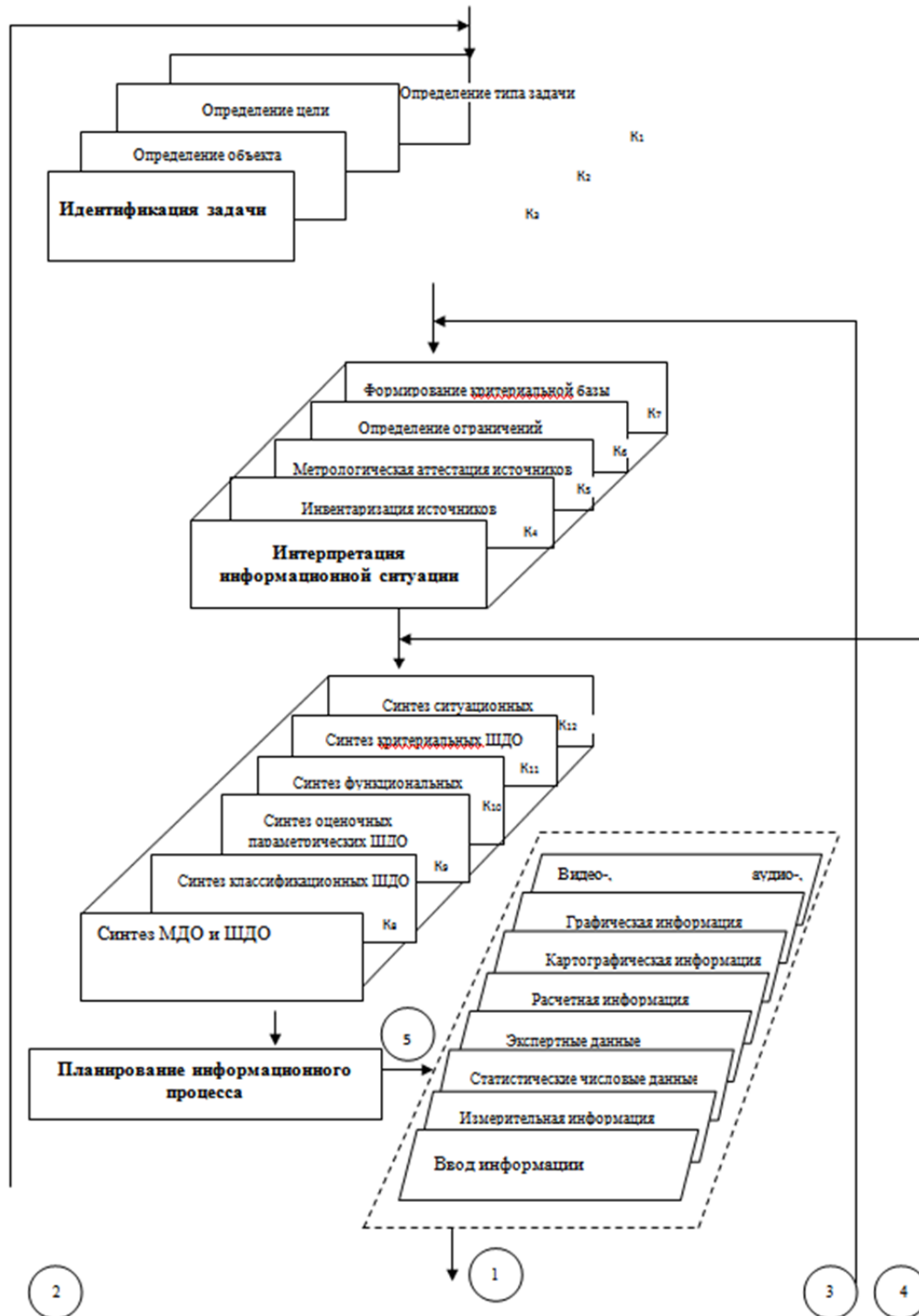
Формулы (36)–(41) иллюстрируют тот факт, что чем меньше наших знаний об объекте и среде его функционирования, тем существеннее степень неопределенности (условности) модели и тем значительнее остающаяся неопределенность в получаемых решениях. Таким образом, если в условиях значительной неопределенности со временем модель не меняется, не развивается с накоплением информации, то неопределенность не снимается, уровень энтропии наших знаний об объекте не снижается, а даже может повышаться за счет того, что некоторые свойства среды уже могут не оказывать влияния на объект, в то время как появятся другие влияющие свойства, но не учтенные в модели, что в свою очередь может привести к снижению качества решений вплоть до абсолютно недостоверных. Поэтому принципиальным является создание методологических и информационно-технологических разработок, обеспечивающих более полное и глубокое «погружение» модели в систему объекта на основе развития структуры модели в смысле возможности изменения составляющих уравнения (37) в направлении снижения остаточной неопределенности.

Процесс целенаправленного развития модели может быть решен на основе принципов управления структурой модели и информационными технологиями на ее основе с позиций общей теории управления в условиях неопределенности.

Так, задача управления моделью объекта и информационными технологиями может быть представлена как некорректная обратная задача, при решении которой должны обеспечиваться условия существования, наблюдаемости, интерпретируемости, квазиустойчивости, квазиединственности (счетное контролируемое множество) решений.

Практическая реализация этих принципов базируется на основе решения задач мониторинга, контроля (аудита) и генерации регулирующих управляющих решений по корректировке модели и технологий. Каждая из этих задач требует своего методологического, информационно-технологического, нормативно-технического, организационного и правового обеспечения. Эти актуальные вопросы требуют детального рассмотрения и изучения для создания действительно мощной произ-

водственной базы разработки современных информационных технологий. На рисунке, приведенном ниже иллюстрируется схема создания и функционирования развивающихся информационных технологий в условиях неопределенности и критериального базиса $\{K_i\}$.



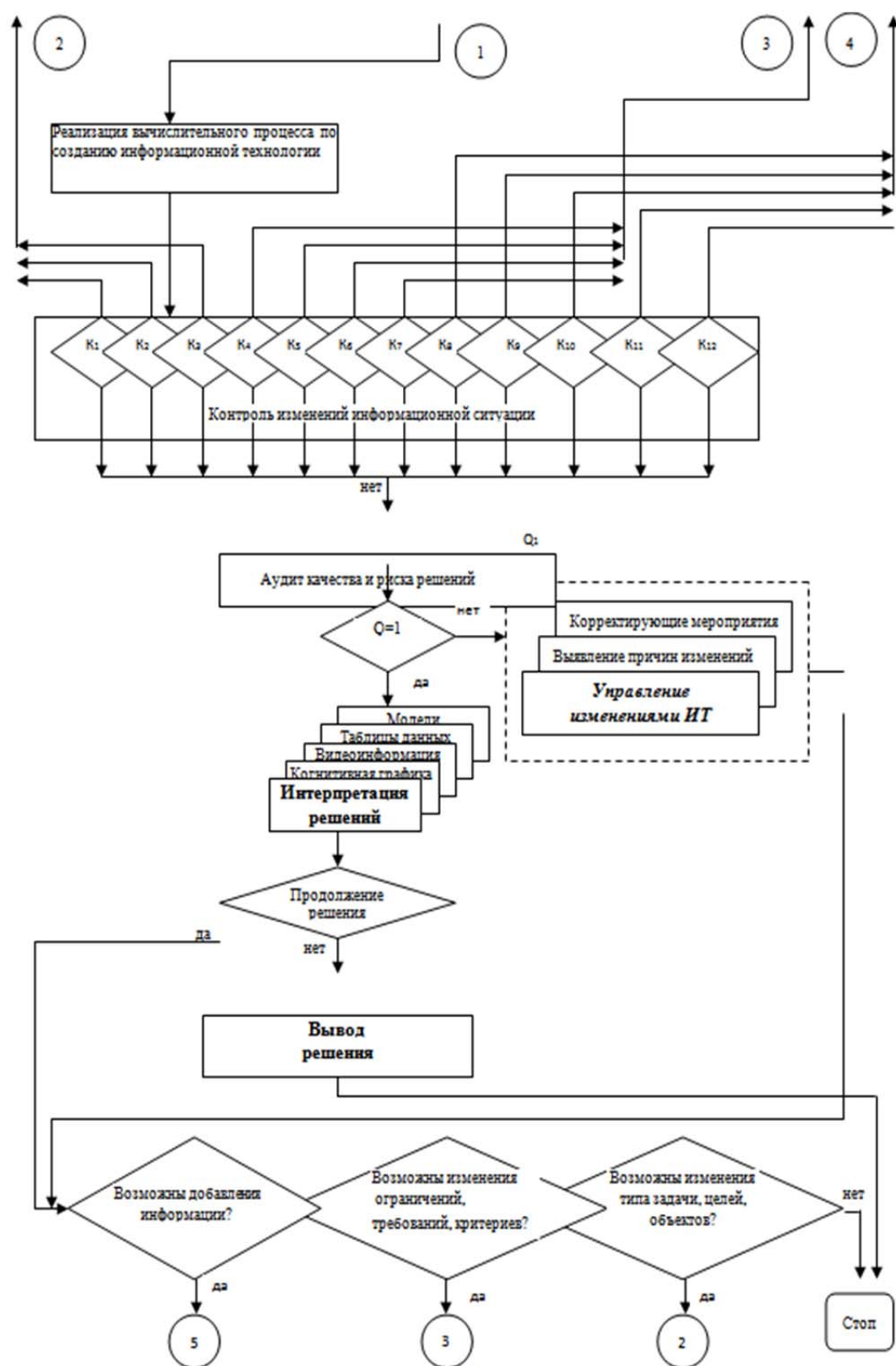


Рис. 9

Рис. 17. Схема создания и функционирования развивающихся информационных технологий в условиях неопределенности и критериального базиса $\{K_i\}$

Таким образом концептуально определена актуальность и принципиальная обоснованность задач создания развивающихся информационных технологий и систем управления ими.

Регуляризирующий байесовский подход (РБП) [1] стал одним из успешных примеров создания платформенных решений для задач моделирования и создания развивающихся информационных технологий, а также задач мониторинга, аудита и управления ими, отражающим указанные выше требования.

Необходимость контроля и обеспечения качества решений обусловило тенденцию сближения методов и технологий искусственного интеллекта (и, как его составляющей, методов мягких вычислений [3]) с методологическими принципами и алгоритмической базой теории измерений. Примером тому является создание в девяностых годах прошлого столетия теории интеллектуальных измерений, байесовских интеллектуальных измерений, а в 1996 году теории мягких измерений [3].

Значительное число разработок искусственного интеллекта использует для представления информации различного рода шкалы [2]. Это объясняется естественностью представления знаний на шкалах, желанием структурировать и классифицировать априорную и апостериорную информацию, а также необходимостью контролировать устойчивость получаемых решений, поскольку подобные задачи относятся к классу некорректных и требуют регуляризации, которая может быть достигнута путем введения шкал.

1.3.3. Интеллектуальные технологии решения многокритериальных задач в условиях риска и неопределенности на основе регуляризирующего байесовского подхода

При принятии оценочных, аудлирующих и управленческих решений необходим инструментарий, позволяющий на основании сформированного запроса в приемлемый интервал времени определять максимально полные, объективные и достоверные решения. Таким образом, требуются системы поддержки принятия решений (СППР), способные к оптимизации выбора решений в многокритериальной среде. Информация, хранящаяся в электронном виде, как правило имеющая статистический, картографический или экспериментальный (измерительный и/или экспертный) вид, характеризуется неточностью, нечеткостью и неполнотой, что определяет условия неопределенности и риска выбора неправильного решения. Кроме того, неопределенность

создается также и за счет действия субъектов сфер деятельности (например, предприятий, банков, торговых организаций в экономике), целенаправленное воздействие которых на ситуацию не может быть заранее определено. Первый тип неопределенности можно считать принципиальным и объективным, второй – принципиальным для сфер активной человеческой деятельности и субъективным. Таким образом, возникает задача принятия решений в многокритериальной среде в условиях риска и неопределенности.

Эффективность выбора решений в информационно-аналитических системах напрямую зависит от оптимизационной мощности и адаптивности к условиям принятия решений математического аппарата, реализованного в них. Рассмотрим схему формирования математической модели при принятии решений. Для этого определим основные компоненты модели принятия решений. Представим объект управления, аудирования или оценивания как сложную динамически развивающуюся систему (СДРС) G_{ot} с набором свойств $\{Q_{oit}\}$ и отношений $\{R_{oit}\}$, среду функционирования или жизнедеятельности объекта как СДРС G_{et} с набором свойств $\{Q_{eit}\}$ и отношений $\{R_{eit}\}$, а также управляющую систему или ЛПР как СДРС G_{at} со свойствами $\{Q_{ait}\}$ и отношениями $\{R_{ait}\}$. Индекс t определяет момент времени для модели СДРС. Индексы i, j в реальности стремятся к бесконечности, а в соответствующих моделях G_{otm} , G_{etm} и G_{atm} они ограничены значениями I_{ot} , I_{et} , I_{at} , которые в каждый момент времени $t=tk$ определяют степень полноты и адекватности моделей. Под объектом оценивания, аудирования или управления можно понимать природные (абиотические и биотические) и антропогенные (техногенные, социально-экономические объекты и процессы), под средой функционирования, обитания или жизнедеятельности природную или техногенную (производственную, социальную, ландшафтную и др. среды), под системой управления (или управляющей, управленческой системой) – техническую, социально-политическую, природную среды или лиц принимающих решения.

Управляющая система (УС) всегда активна, действует целенаправленно в целях перевода объекта управления (ОУ) в требуемое состояние. В качестве инструмента воздействия УС имеет в своем архиве или генерирует распоряжения, рекомендации, выводы, стратегии и тактики управления. Модель управляющей системы G_{ami} , реализуемая средствами информационных технологий, не может быть полностью

адекватна реальной УС, так как она постоянно активно развивается вместе с ее приоритетами и предпочтениями в каждой конкретной ситуации. Кроме того, в общем рассмотрении, среда может существенно влиять на изменение состояния УС, также как и активный ОУ. Причем это влияние может быть случайным, неорганизованным или целенаправленным в соответствии с интересами среды и ОУ.

Среда жизнедеятельности или функционирования объекта может быть активной или пассивной в зависимости от типа задачи, но всегда остается динамичной и недоопределенной в своем модельном представлении и, следовательно, не может иметь устойчивого модельного представления на детализированном уровне, необходимом для предварительной формализации процессов и факторов влияния среды на ОУ и УС.

Объект управления, как уже рассматривалось, является в общем случае СДРС и может быть активным или пассивным по отношению в среде и УС. Таким образом, в сложных задачах, в отличие от предположения, выдвинутого в [2], адекватную детализированную математическую модель ОУ определить нельзя.

Однако для практических задач можно установить уровень детализации, который позволит создать в рамках конкретных ограничений, требований и набора критериев *квазиадекватные математические модели* ОУ, среды, УС, устойчивые в пределах перечисленных составляющих процесса принятия решений. При этом необходимо определить метрологические характеристики моделей и решений, устанавливающие степень адекватности, полноты и достоверности моделей.

В условиях вышеуказанной неопределенности методологической основой для создания моделей СДРС может служить *регуляризирующий байесовский подход (РБП)*. Рассмотрим процесс принятия решений и его формализацию на основе методологии РБП и байесовских интеллектуальных технологий [3–5], которые реализуют РБП в различных типовых информационно-аналитических задачах.

Для формализации модельного представления процесса принятия решений определим основные понятия, в частности, понятия функции реализации f_t , платежной функции P_t и целевой функции C_t .

Согласно [2] под *функцией реализации f_t* понимается такая функция, что:

$$f_t: G_{omt} * G_{emt} * G_{amt} * Y_t \rightarrow \{h_{kt} | \{MX\}_{kt}\}, \quad (42)$$

где $\{h_{kt} | \{MX\}_{kt}\}$ – множество альтернативных решений с комплексами метрологических характеристик, содержащих показатели точности, надежности и достоверности решений h_{kt} , принадлежащих динамическому компактному решению H_t

Определения и допустимые функциональные преобразования $h_{kt}, \{MX\}_{kt}, H_t$ приведены в [6].

В уравнении (42) предполагается, что функция реализации f_t , принадлежащая функциональному множеству F_t способна адаптироваться к изменениям условий принятия решений и менять свою структуру во временном диапазоне функционирования СППР.

Поскольку при реализации процесса принятия решений на методологической основе РБП все преобразования производятся на шкалах ШДО, которым соответствуют СДРС обозначаемые как $S_{omt}, S_{emt}, S_{amt}, S_{yt}$ соответственно, то для БИТ формула (42) может быть представлена в следующем виде:

$$f_{st}: S_{omt} * S_{emt} * S_{amt} * S_{yt} \rightarrow \{h_{skt} | \{MX\}_{skt}\}, \quad (43)$$

где $\{h_{skt} | \{MX\}_{skt}\}$ – множество альтернативных решений в виде регуляризированных байесовских оценок (РБО.), представляющих собой реперы ШДО. Свойства РБО. определены в [6]. Как показано в [1–7], РБО могут представлять собой все типы информационных потоков: числовой лингвистической, графической. Картографической и других видов информации. По типу информации ШДО разделяются на соответствующее число типов. Числовые и лингвистические ШДО могут быть представлены в виде сопряженной ШДО, основного типа носителя шкалы для БИТ. Лингвистические ШДО относятся к порядковым шкалам, а числовые представляют собой шкалы отношений.

Вид сопряженной шкалы представлен на рисунке 18.

Уравнение (43) представляет собой **концептуальную математическую модель реализационной структуры СППР на основе РБП**. Вторая составляющая математической модели принятия решений на основе РБП представляет собой оценочную структуру по определению данному в 2 или в терминологии РБП платежную матрицу Π_t (матрицу штрафов). Вид платежной матрицы, как правило, выбирается заранее на основании расчетных или субъективных

подходов ЛПР. Подробнее вопрос выбора матрицы Π_t обсуждается в [6]. Оценочная функция определяет выигрыш (прибыль) УС при реализации эффективного решения или плату за неверное решение, причем вид матрицы позволяет определить плату или выигрыш при конкретном сочетании решений и состояния основных компонентов функции реализации.

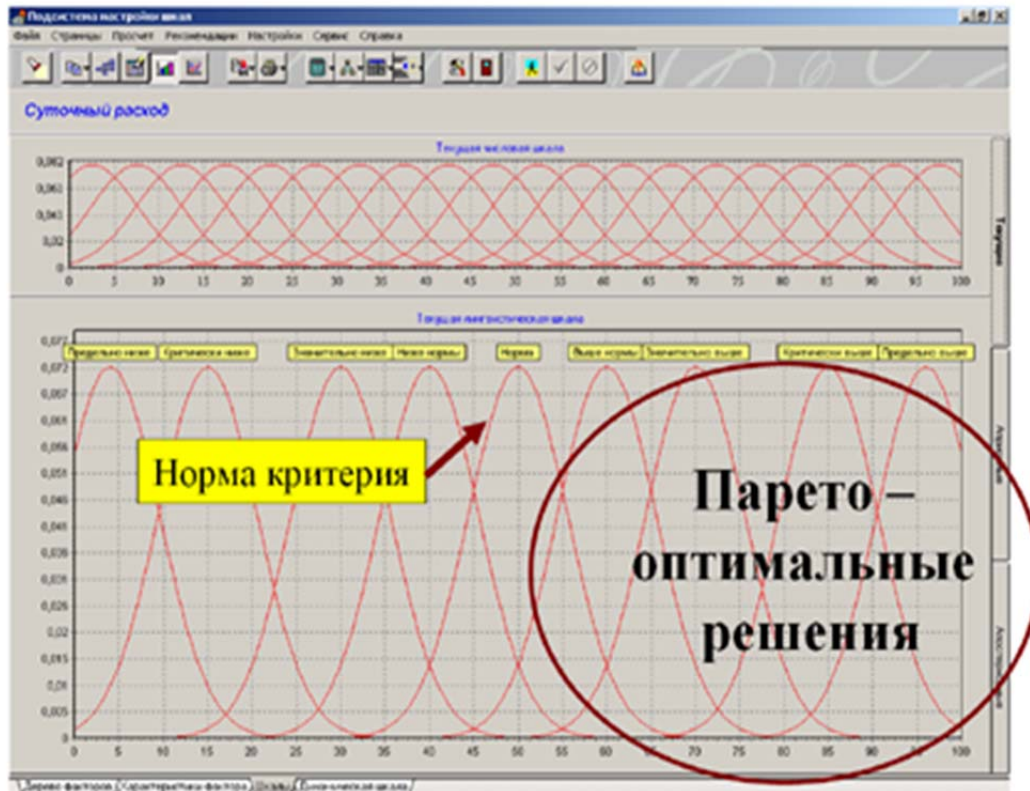


Рис. 18.

Задача СППР состоит в выборе оптимальных альтернатив $\{h_{skt} | \{MX\}_{skt}\}$, удовлетворяющим условиям оптимума целевой функции z_t , принадлежащей множеству Z_t .

Целевая функция основывается на критериальной базе C_t , где C_t имеет вид многомерного критерия, состоящего совокупности из разнотипных критериев $\{c_{it}\}$. Далее будет проиллюстрирован механизм принятия решений на ШДО в многокритериальной среде C_t .

Целевая функция z_t представляет собой свертку функций реализации f_t и оценивания Π_t , что может быть записано в виде:

$$z_t = \Pi_t * f_t \quad (44)$$

В качестве альтернативных решений выбираются решения, доставляющие экстремальные значения целевой функции по обобщенному критерию C_t или отдельным критериям c_t .

$$\{h_{skt} | \{MX\}_{skt}\} = \text{argextr} \{c_{it}(\Pi_t * f_t)\} \quad (45)$$

Так как аппарат выбора решений на основе РБП всегда базируется на основных принципах и критериях классического правила выбора гипотез Байеса, то основополагающим критерием c_{bt} всегда будет минимум среднего риска решения (максимум апостериорной достоверности решения с учетом вида платежной матрицы).

С учетом этого формулу (45) можно переписать в виде:

$$\{h_{skt} | \{MX\}_{skt}\} = \text{argextr} \{c_{it}(\text{argextr} \{c_{bt}(\Pi_t * f_t)\})\} \quad (46)$$

Таким образом, в соответствии с формулой (46) *математическую модель принятия решений на многокритериальной РБП основе* будет выбираться ряд наиболее значимых альтернативных решений с комплексами характеристик их качества и риска.

Рисунок 19 иллюстрирует выбор альтернативных решений на многокритериальной РБП основе в СППР, реализующей инструментарий БИТ для одного из проектов управления ресурсами региона.

Процесс принятия решений на основе БИТ в многокритериальной среде можно представить в виде композиционной свертки РБО оптимальных по каждому из критериев c_{it} и полученных на сопряженных ШДО. При этом на лингвистических шкалах сопряженных ШДО устанавливается значение критерия c_{it} в виде «нормы». Так как ШДО имеет характер порядковой шкалы, то они обладают отношением доминирования по Парето [6, 7]. Очевидно, все решения в виде РБО, лежащие выше (правее) репера «нормы» будут Парето-оптимальными с точки зрения соответствия конкретному критерию данной шкалы (см. рисунок 20). Такая шкала создается для каждого критерия множества $\{c_{it}\}$. Свертка решений РБО для всех критериев дает интегральное Парето-оптимальное решение, доставляющее экстремум уравнению (46).

Таким образом, Парето-оптимальные решения являются подмножеством множества H_t и могут быть представлены в виде РБО на всем пространстве решений компакта H_t . При этом отпадает необходимость сужения области Парето, так как в состав оптимальных в смысле (46) РБО входят только значимые по достоверности альтернативы.

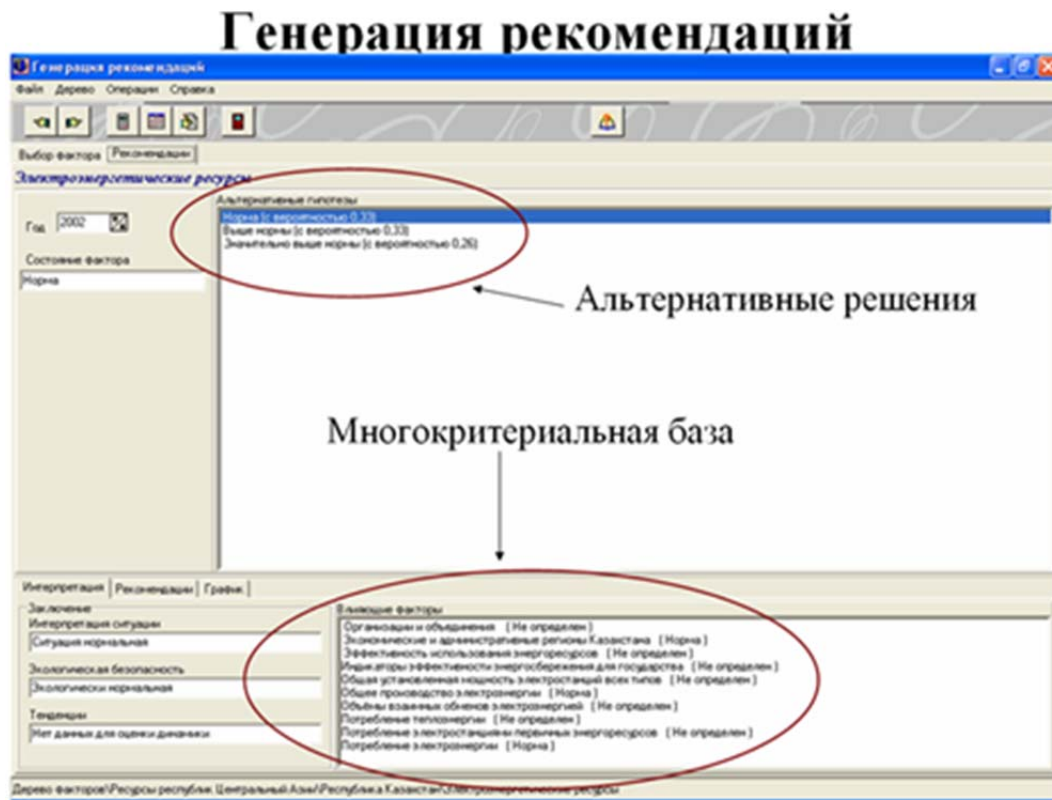


Рис. 19.

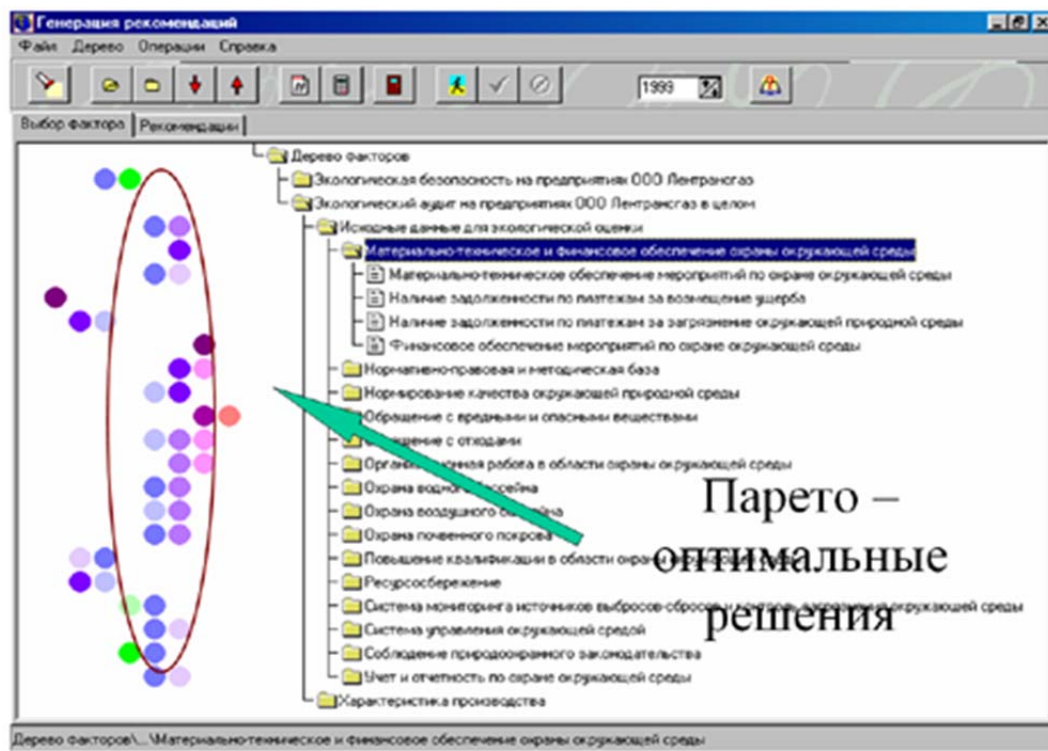


Рис. 20.

Рисунок 20 иллюстрирует так называемую таблицу (карту) риска решений. Кружками отмечены значимые альтернативы, а их цветовое соотношение показывает степень риска от незначительного (синий) до опасного (красный). Оценка «Норма» или точное соответствие критерию интерпретируется зеленым цветом. РБО в виде нескольких альтернатив свидетельствует о степени нечеткости решения из-за исходной неопределенности информации и ситуации принятия решений.

Совокупность РБО соответствующая совокупности значений критериев «норма» определяет на гиперкубе ШДО модель образцовой, стандартной ситуации или состояния объекта, с которой можно сравнивать их текущее состояние, определенное по экспериментальным данным. И поскольку ШДО метризованы, и на них определены мера и метрика расстояния между реперами, то можно при проведении аудита состояния определить не только отличие его от образцового, но и указать, насколько оно далеко отстоит от образцового в единицах принятой метрики расстояния. Таким же образом определяется модель наиболее благоприятного и критического состояний объекта или ситуации.

В задачах многокритериального выбора вводят понятие важности критерия W_{cit} для установления его приоритета перед остальными критериями. Если имеется всего J критериев, то в многокритериальных БИТ вес (приоритет) критерия можно установить по формуле:

$$W_{it} = \sum_{i=1}^I \sum_{j=1}^J \Pi_{ij} \quad (47)$$

В соответствии с формулой (47) определяется матрица важности или приоритетов критериев в виде W_{ijt} .

Математическая модель принятия решений на основе БИТ в многокритериальной среде может быть с учетом (47) представлена в виде:

$$\{h_{skt} | \{MX\}_{skt}\} == \text{argextr} \{c_{it}(\text{argextr} \{c_{bt}(W_t * \Pi_t * f_t)\})\} \quad (48)$$

Процесс принятия решений на основе согласно [1] может быть представлен последовательностью следующих этапов:

1. *Подготовительный этап*, к подэтапам которого относятся:
 - Постановка и концептуальная формализация задачи принятия решений.

- Определение требований к качеству и ограничений.
- Формирование критериальной базы (важности и степени приоритетности, предпочтительности критериев и их взаимосвязи).
- Инвентаризация информационных потоков.
- Метрологическое обоснование информационных потоков.
- Инвентаризация знаний о решениях, создание базы знаний по реализованным рекомендациям, опыту применения возможных решений и реализации сформированных стратегий.
- Выбор реализационной функции.
- Выбор платежной (оценочной функции).
- Формализация оптимизационного уравнения.

2. Этап выбора решения:

- Получение значений показателей, параметров, индикаторов состояний объектов, среды и управляющей системы, которая при активном режиме может выступать в качестве подходе.
- Метрологическое обоснование качества решений
- Расчет платежей и экономических показателей при реализации решений
- Формирование списка альтернатив на основе критериев, требований и ограничений, а также априорного опыта

3. Этап реализации решений.

4. Этап формирования и проигрывания сценариев.

5. Этап проверки качеств альтернативных решений:

- Получение новых значений показателей.
- Получение новых оценок ситуации.
- Формирование списка альтернатив.
- Принятие решений о готовности и годности альтернатив, а иначе возвращение ко второму этапу.

Обобщенная оптимизационная модель выбора решений на основе БИТ может быть представлена как композиция (свертка) компонентов оптимизационных уравнений вида (48) для каждого этапа и подэтапов указанных выше в виде:

Например, для 5 указанных этапов уравнение имеет вид:

$$\{h_{skt} | \{MX\}_{skt}\} = \text{argextr} \{c_{it}(\text{argextr} \{c_{bt}(\sum_{i=1}^5 W_{it} \sum_{i=1}^5 \Pi_{it} \sum_{i=1}^5 f_{it})\})\} \quad (49)$$

Альтернативные решения в виде нечеткой РБО с учетом важности критерия для каждого из подэтапов можно получить по модифицированной формуле Байеса:

$$h_{klt} = \sum_{i=1}^I \sum_{j=1}^J W_{ilt} * P_{ilt}(h_{ilt}) * W_{jet} * P_{jet}(h_{jet}) \quad (50)$$

Модель (48) принятия решений на основе БИТ в многокритериальной среде предназначена для решения многокритериальных задач в условиях риска и неопределенности, а также активности и динамичности объекта управления среды и управляющей системы, что в концептуальной постановке охватывает широкий круг сложных практических задач.

1.3.4. Методология и технологии мягких критериев и мягкого нормирования

Мягкими критериями назовем критерии, выраженные в вероятностном, нечетком, многоальтернативном виде совокупности критериев с соответствующей степенью их вероятности.

Практика работы со сложными системами технического и социально-экономического типов определила потребность в разработке **методов «мягкого нормирования»**. Так в работе [4] проведен подробный аналитический обзор системы нормирования труда при проектировании и изготовлении сложных технических систем на примере современных корабельных систем. В монографии на основании изучения потребностей практики создания новых

образцов техники сделан вывод о необходимости разработки новых подходов к нормированию показателей производственных процессов и оценке трудоемкости в условиях неопределенности и нестабильности экономики.

В связи с тем, что нестабильность ситуаций не позволяет заранее установить жесткие критерии и нормативы как единственно возможные (что было в плановой экономике), такие подходы и методы нормирования в отличие от существующих должны обеспечивать проверку на соответствие мягким критериям, нормирование и аудирование в условиях допусков, многоальтернативности значений критериев. На практике это позволит, например, устанавливать обоснованные сроки и цены изготовления новых изделий повышенной сложности,

наукоемкости и трудоемкости, что, в свою очередь позволит избежать непланируемых задержек выпуска продукции.

На **рисунке 21** представлен вид «мягкого» критерия проверки выборки по Фишеру.

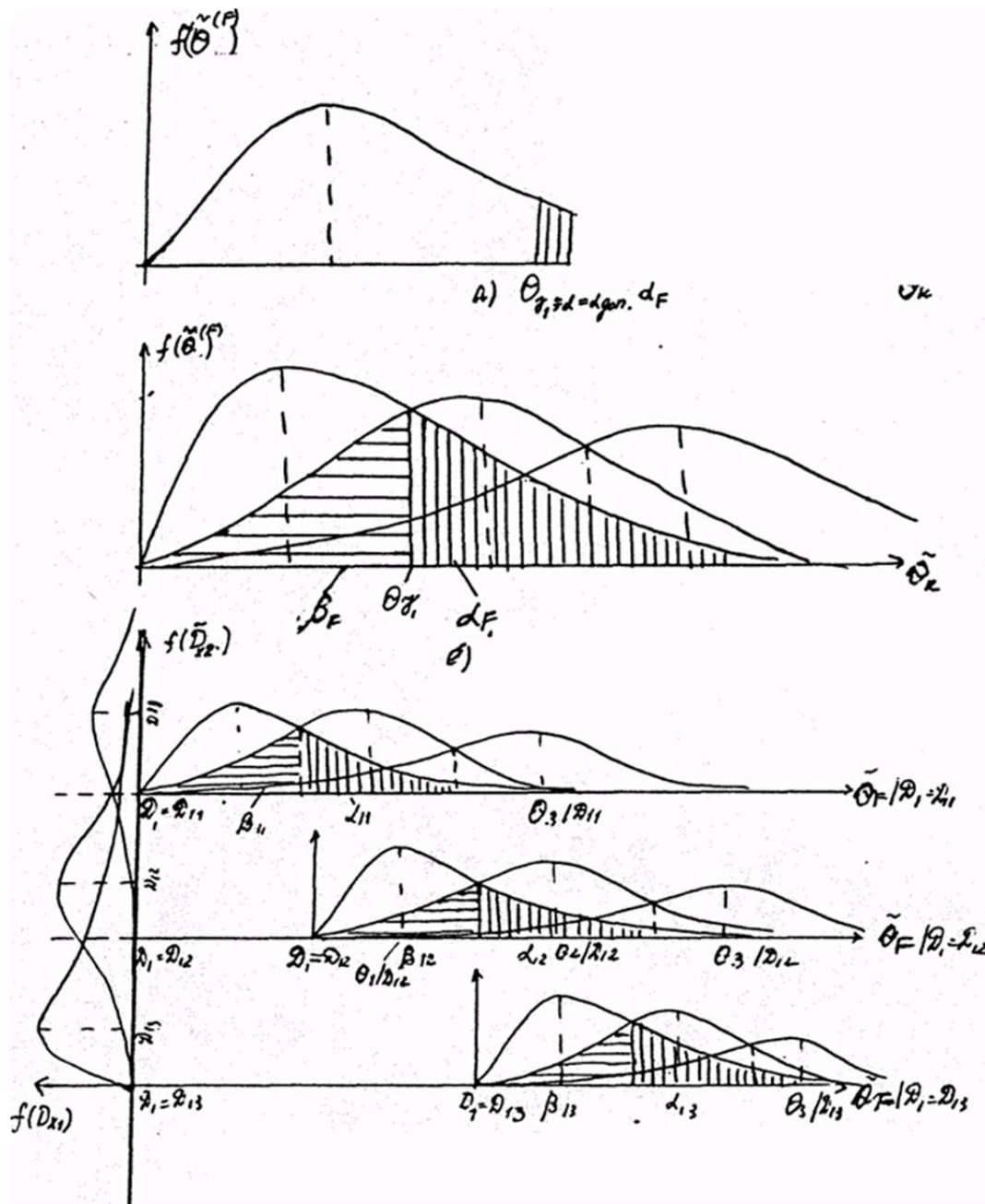


Рис. 21.

Шкала значений критерия представляет собой трехмерную функциональную шкалу, соответствующую трем параметрам нецентральной

ного распределения Фишера (двум параметрам формы и параметру нецентральности). Параметр нецентральности играет в мягких критериях проверки статистических гипотез ключевую роль. Он позволяет проверять многоальтернативные гипотезы и контролировать уровень ошибок 2 рода, что в классическом виде критериев отсутствует.

Можно отметить, что в целом такой вид критерия соответствует типу «мягких» функций, рассмотренных в статье ниже.

Моделирование на основе мягких подходов

Как показано в ряде работ, [1–5] такие подходы реализуются в рамках концепций построения самоорганизующихся моделей и методов.

В условиях значительной неопределенности модель сложного объекта и среды его окружения должна изменяться в зависимости от получаемой информации и меняющихся требований, ограничений, целевых функций и критериев выполняемой задачи.

Технологии моделирования при этом не имеют жесткой алгоритмической структуры и могут развиваться автономно или пользователем по мере изменения требований, ограничений или целей решаемых задач. Кроме того, получаемые решения носят характер *«мягких», то есть вероятностных, нечетких, многоальтернативных оценок и моделей*. Такие технологии не имеют жесткой алгоритмической структуры и могут развиваться автономно или пользователем по мере изменения требований, ограничений или целей решаемых задач. Кроме того, получаемые решения носят характер *«мягких», то есть вероятностных, нечетких, многоальтернативных оценок и моделей*.

Процесс целенаправленного развития модели может быть решен на основе принципов управления структурой модели и информационными технологиями на ее основе с позиций общей теории управления в условиях неопределенности.

Так, задача управления моделью объекта и информационными технологиями может быть представлена как некорректная обратная задача, при решении которой должны обеспечиваться условия существования, наблюдаемости, интерпретируемости, квази-устойчивости, квази-единственности (счетное контролируемое множество) решений.

Практическая реализация этих принципов базируется на основе решения задач мониторинга, контроля (аудита) и генерации регулирующих управляющих решений по корректировке модели и технологий. Каждая из этих задач требует своего методологического, инфор-

мационно-технологического, нормативно-технического, организационного и правового обеспечения. Процесс целенаправленного развития модели может быть решен на основе принципов управления структурой модели и информационными технологиями на ее основе с позиций общей теории управления в условиях неопределенности.

Так, задача управления моделью объекта и информационными технологиями может быть представлена как некорректная обратная задача, при решении которой должны обеспечиваться условия существования, наблюдаемости, интерпретируемости, квази-устойчивости, квази-единственности (счетное контролируемое множество) решений.

Практическая реализация этих принципов базируется на основе решения задач мониторинга, контроля (аудита) и генерации регулирующих управляющих решений по корректировке модели и технологий. Каждая из этих задач требует своего методологического, информационно-технологического, нормативно-технического, организационного и правового обеспечения.

Регуляризирующий байесовский подход (РБП) [1] стал одним из успешных примеров создания платформенных решений для задач моделирования и создания развивающихся информационных технологий, а также задач мониторинга, аудита и управления ими, отражающим указанные выше требования. Необходимость контроля и обеспечения качества решений обусловило тенденцию сближения методов и технологий искусственного интеллекта (и, как его составляющей, методов мягких вычислений [3]) с методологическими принципами и алгоритмической базой теории измерений. Примером тому является создание в девяностых годах прошлого столетия теории интеллектуальных измерений, байесовских интеллектуальных измерений, а в 1996 году теории мягких измерений [3].

1.3.5. Мягкие функции и их применение в управлении сложными системами

Для повышения эффективности управления системами необходимо обеспечивать высокую степень формализации описания структуры, функций и свойств систем.

Формализованные модели также имеют вид «мягких» функциональных зависимостей в указанном выше смысле. Назовем их *«мягкими функциями»*.

Методы определения мягких функций должны иметь вполне определенную последовательность этапов решения задач создания и

развития информационных технологий обработки информационных потоков данных и знаний с целью их формализации.

Такая последовательность может быть следующей:

Этап 1. Разработка концептуальной модели сложной системы в виде иерархической структуры факторов (собственных и внешней среды), в условиях имеющейся априорной информации, критериев и метрологических ограничений и требований

Этап 2. Сбор информации о факторах модели и их взаимосвязях.

Этап 3. Построение регуляризирующих информационных технологий оценки и измерения факторов.

Этап 4. Разработка методов контроля (аудита) факторов на соответствие установленным или определяемым в ходе решения задачи нормам и критериям.

Этап 5. Определение формализованного вида функциональных зависимостей для описания состояний, свойств, ситуаций сложных систем в их непрерывном развитии и взаимодействии.

Этап 6. Проверка адекватности полученных моделей, их соответствия имеющимся данным.

Этап 7. Интеграция полученных моделей в процесс обработки информации.

Этап 8. При реализации функций управления сложными системами необходима генерация *мягких управленческих решений, определенных в концепции мягкого управления в.[5] в виде многоальтернативного управленческого решения, представляющего собой совокупность управленческих решений с различными вероятностями (возможностями) его альтернатив*

Этап 9. Выбор и реализация в процессе управления управленческого решения в единственного четкого решения из предложенных выше альтернатив.

Всем указанным этапам должны соответствовать методы и информационные технологии, имеющие «мягкие» подходы в своей основе.

На **рисунках 22 и 23** приведены результаты аппроксимации результатов, полученных методами РБП в виде «мягких» функций авторегрессии третьего порядка и апостериорных распределений оценок типовыми законами распределений из классов системы Пирсона, для чего используется математическая библиотека РБП.

Полученные модели законов распределений используются затем для интерпретации оцениваемых свойств объекта, для генерации выборок большего объема и последующей ее обработки мощными мето-

дами математической статистики, для управления качеством получаемых решений и их риском.

Для оценки параметров и показателей применяются статические и динамические параметрические оценочные шкалы с динамическими ограничениями (ШДО), вид которых представлен на [рис. 21, 22, 23](#). Для динамических показателей реализованы динамические ШДО (см. [рис. 22](#)). Причем, как первый, так и второй тип шкал реализован виде сопряженных ШДО для обработки и хранения как числовой, так и лингвистической информации.

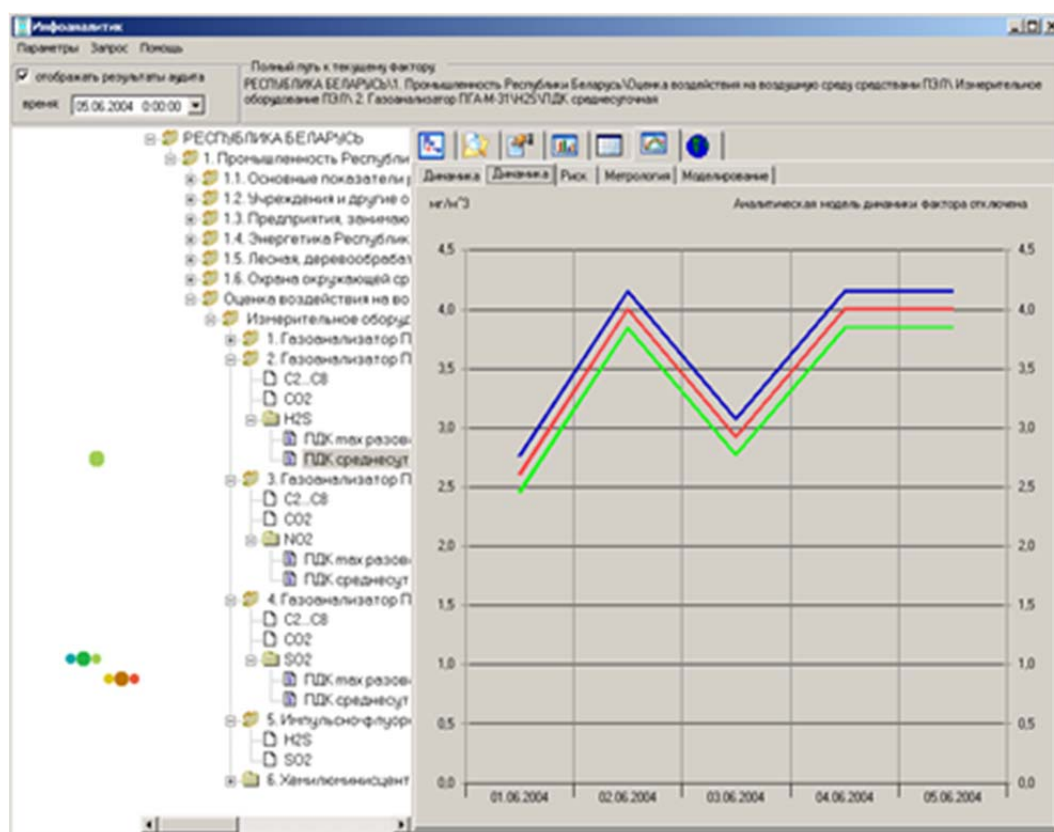


Рис. 22. «Мягкая» функция авторегрессии

Числовые ШДО имеют тип шкал отношений и обладают богатыми вычислительными возможностями. В зависимости от изменения характеристик информационных потоков меняется структура этих шкал.

Для функциональных преобразований созданы функциональные ШДО, теория построения которых приведена в [1].

Задачи нормирования, критериального выбора и аудита производятся на критериальных шкалах. Вид такой шкалы приведен на рисунке 4. При смене критериальной базы или изменении свойств кон-

тролируемого объекта, когда меняется норма показателей его свойств, шкала настраивается на новые значения норм автоматически или пользователем. Таким образом может контролироваться динамика норм и критериев. Кроме того, эти шкалы могут быть сопряженными для реализации контроля одного и того же показателя по множеству различных критериев.

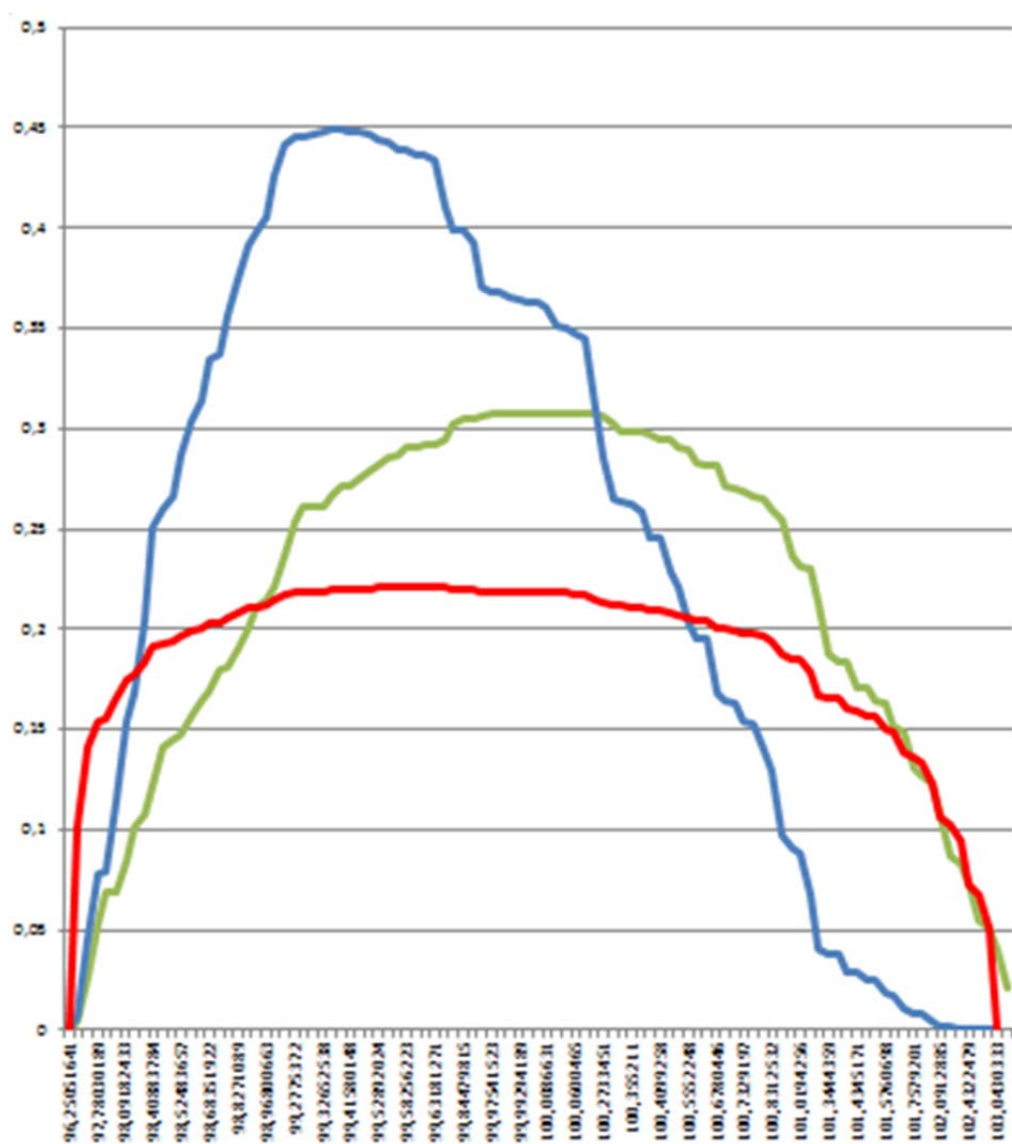


Рис. 23. «Мягкая» функция плотности распределения

Критериальные шкалы могут быть полезны в задачах контроля (экологического, энергетического, технического и других) для установления предупреждающих границ при вероятностном представле-

нии контролируемого показателя. Например, существует практически важная задача оценки вероятности превышения значений контролируемого параметра допустимого предела и в соответствии с этим изменения предупреждающих границ контроля. На основе БИТ на функциональных ШДО в режиме мониторинга определяется изменение закона распределения параметра (см. рис. 7), например, в следствии изменения технологического оборудования, режимов производства, материала заготовок. По результатам оценки закона распределения параметра принимаются решения об изменении предупреждающего значения на критериальной шкале БИТ, что предупреждает появление брака, экологических или энергетических рисков. Подробнее механизм реализации этого подхода рассмотрен в [3].

1.3.6. Методология и технологии байесовской математической статистики на основе РБП. Система математической обработки статистической информации «байесовская математическая статистика»

Мягкие статистические измерения на основе регуляризирующего байесовского подхода.

Методологии РБП и БИИ стали основой для разработки методов и средств регуляризированной байесовской статистики, сочетающей свойства классических методов математической статистики с информационными технологиями БИИ.

Это направление развития методов вероятностного подхода возникло в ситуации, когда применение методов классической статистики в прикладных областях в сложных информационных ситуациях (малых выборок экспериментальных данных, неполноты факторов, недостоверности, нерегулярности и субъективности данных) оказалось крайне неэффективным.

Тому есть объяснения теоретического, социально-экономического и технического уровней рассмотрения.

В условиях сложных информационных ситуаций, как правило, теоретические допущения и ограничения не соответствовали реальным, соответствующим функционирующим в действительности объектам.

Нарушение плановости промышленных производств в последние годы сказалось на качестве и объеме статистической информации, поступающей обычно от ведомственных организаций в качестве статотчетности.

Уникальность природных и вписанных в природную среду техногенных объектов, их изменчивость не давали оснований считать наблюдения, полученные в ходе их изучения в течение какого-либо периода времени повторяющимися экспериментами, допускающими их частотную интерпретацию. В этих условиях приходилось достаточно часто при применении методов классической статистики принимать необоснованные субъективные решения относительно вероятностных характеристик свойств реальных объектов, в частности, закона распределения, что влияло на качество результатов. Имеющаяся дополнительная априорная информация, представленная не в числовой форме, не используется.

Кроме того, современные статистические пакеты обработки экспериментальных данных имеют неудобный жесткий набор алгоритмов оценивания отдельных числовых вероятностных характеристик, не позволяющий достаточно просто для не математика построить информационные технологии, ориентированные на специфику прикладной задачи и имеющейся информации пользователя. Система байесовской математической статистики свободна от этих недостатков.

«Мягкими» статистическими измерениями называются измерения, результатами которых являются мягкие, нечеткие оценки статистических характеристик в виде альтернативных решений с соответствующими им вероятностями.

На основе мягких статистических измерений построены алгоритмы определения всех числовых случайных величин и процессов.

Кроме того система имеет возможность функциональных характеристик, таких как плотность вероятности, корреляционные функции, критерии проверки гипотез, характеристик периодичностей, трендов, взаимовлияния, предполагая участие в процессе оценки самого пользователя, а также возможности учета априорной информации и получения результата на соответствующих шкалах БИИ многоальтернативных решений в виде нечеткого числа, нечеткой функции, нечеткого высказывания.

Методологическая основа БИИ, реализованная в системе, значительно расширяет функциональные и интерпретационные возможности. Так, например, для критериев проверки гипотез, возможно, не только отвергнуть или принять существующую гипотезу, но и указать расстояние от нее до нулевой и ее примерный вид. Особенно важные для практики возможности контроля качества статистических оценок

решений, а также построения оптимального плана статистического эксперимента, обеспечивающего требуемое качество решения поддерживаются информационными технологиями БИИ на всех этапах получения результатов.

Система стала математической базой для ряда прикладных систем мониторинга и поддержки принятия решений в охране окружающей среды и природопользовании.

Общее описание

Программное обеспечение «Байесовская Математическая Статистика» (далее БМС) представляет собой систему для выполнения различных математических операций, вычислений, преобразований и хранения результатов данных операций в базе данных.

БМС представляет собой среду, содержащую алгоритмы для различных вычислений, с возможностью их добавления и удаления.

Система позволяет хранить в себе входные и выходные данные вычислений, составные алгоритмы.

Физически данные хранятся в базе данных MS SQL Server. Имеется возможность сохранения данных и составных алгоритмов в файл для экспорта на другие экземпляры системы.

Также имеется возможность работы различных экземпляров системы с одной сетевой базой данных. Таким образом, реализуется механизм распределенного доступа и обработки данных.

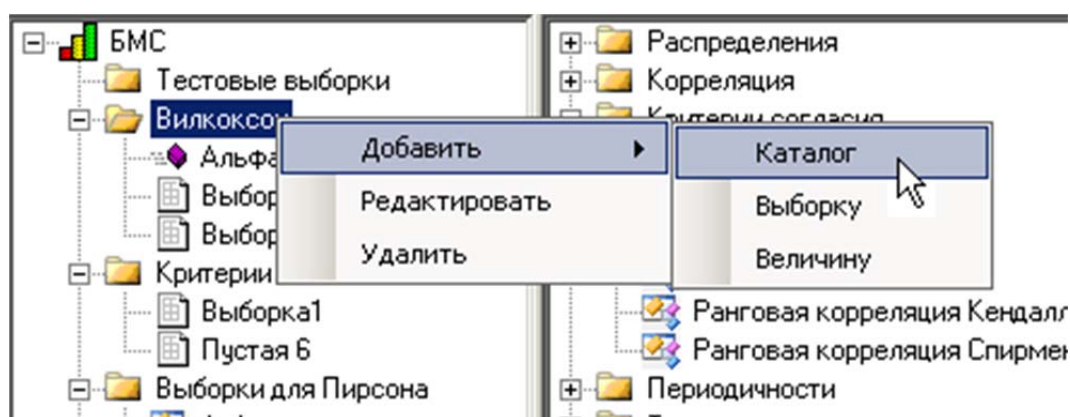
Каждый алгоритм представляет собой динамически загружаемую библиотеку DLL. Составной алгоритм представляет собой логически связанные в сложный алгоритм динамические библиотеки и позволяет расширить функциональность системы.

Подсистема управления данными

Хранение данных реализовано в виде иерархии объектов реализованной для простоты аналогично файловой системе ОС «MS Windows».

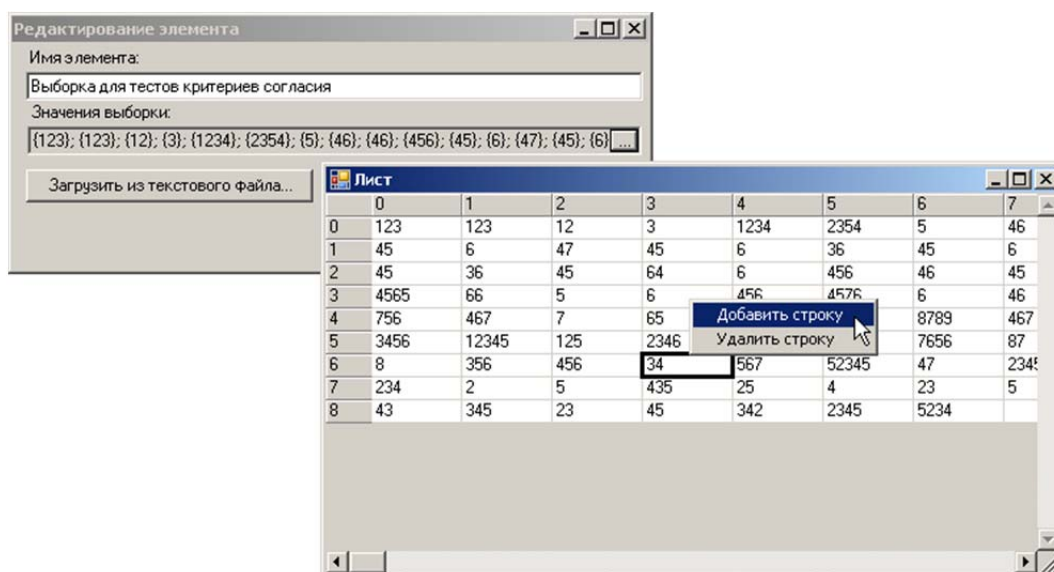
Основными типами данных в системе являются значение и выборка (набор значений или одномерный массив). При функционировании алгоритма несколько выборок могут преобразовываться в многомерные массивы со срезами различной размерности.

Дерево данных состоит из подкаталогов, вложенность которых не ограничена. Манипуляции с деревом данных осуществляются посредством меню и технологии «Drag & Drop», таким образом, упрощая работу пользователя.



Существует возможность загрузки и сохранения данных из файлов определённого формата.

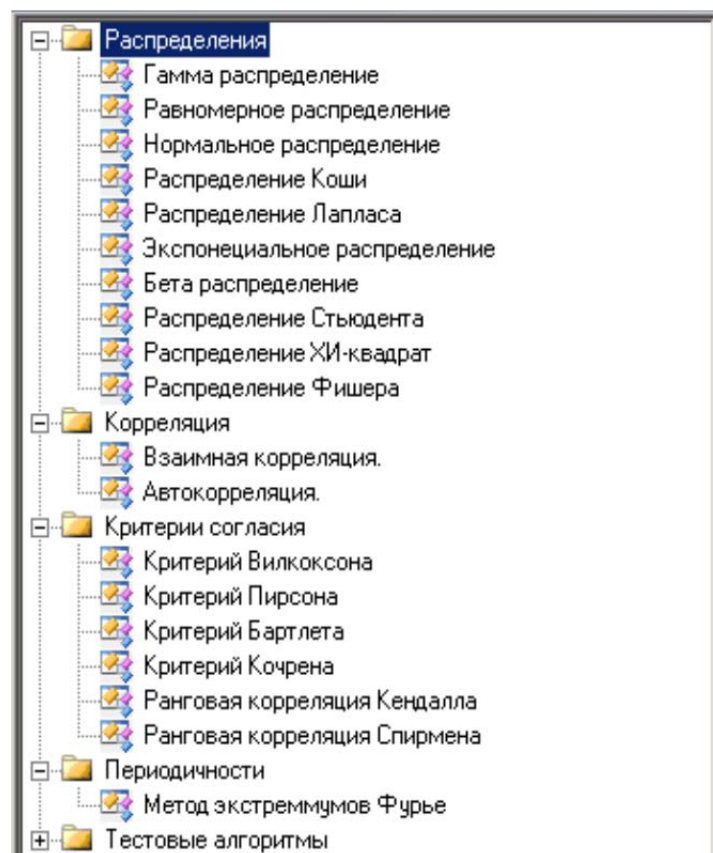
Таким образом, реализована возможность импорта и экспорта данных между различными экземплярами системы.



Ниже приведен список реализованных на данный момент алгоритмов:

- Вычисление характеристик следующих распределений:
 - Гамма распределение.
 - Равномерное распределение.
 - Нормальное распределение.
 - Распределение Коши.
 - Распределение Лапласа.
 - Экспоненциальное распределение.
 - Бета распределение.

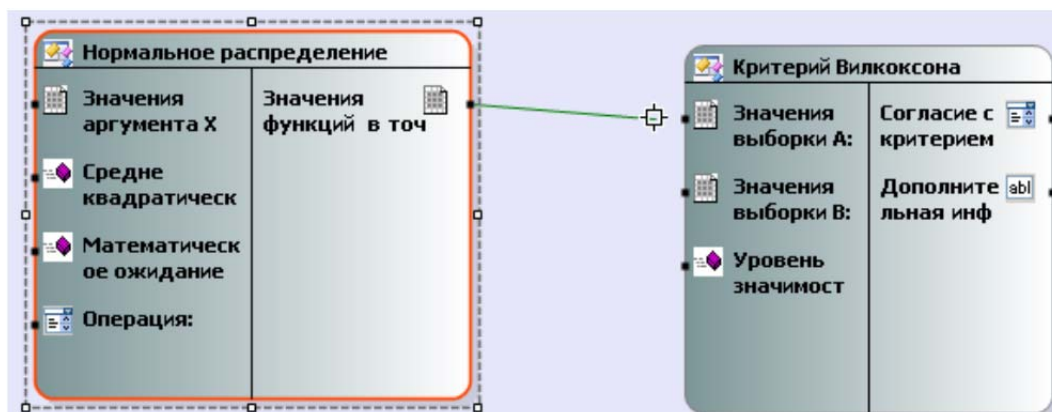
- *Распределение Стьюдента.*
- *Распределение ХИ-квадрат.*
- *Распределение Фишера.*
- Проверка критериев согласия:
 - *Критерий Вилкоксона.*
 - *Критерий Пирсона.*
 - *Критерий Бартлета.*
 - *Критерий Кочрена.*
 - *Критерий Кендалла.*
 - *Критерий Спирмена.*
- Вычисление характеристик выборки:
 - *Авто и взаимной корреляции.*
 - *Периодичности методом экстремумов Фурье.*
 - *Оценки характеристик выборки (Математическое ожидание, дисперсия, асимметрия, эксцесс).*
 - *Составные алгоритмы.*



Сложный алгоритм представляет собой алгоритм созданный из простых содержащихся в управляемых в сборках. Соединяя выходные

значения одних базовых алгоритмов со входными значениями других получаем цепочку выполнения алгоритма. Данный механизм позволяет создавать более функциональные алгоритмы и автоматизировать процесс последовательных рутинных вычислений.

Создание сложного алгоритма происходит с помощью редактора сложных алгоритмов.



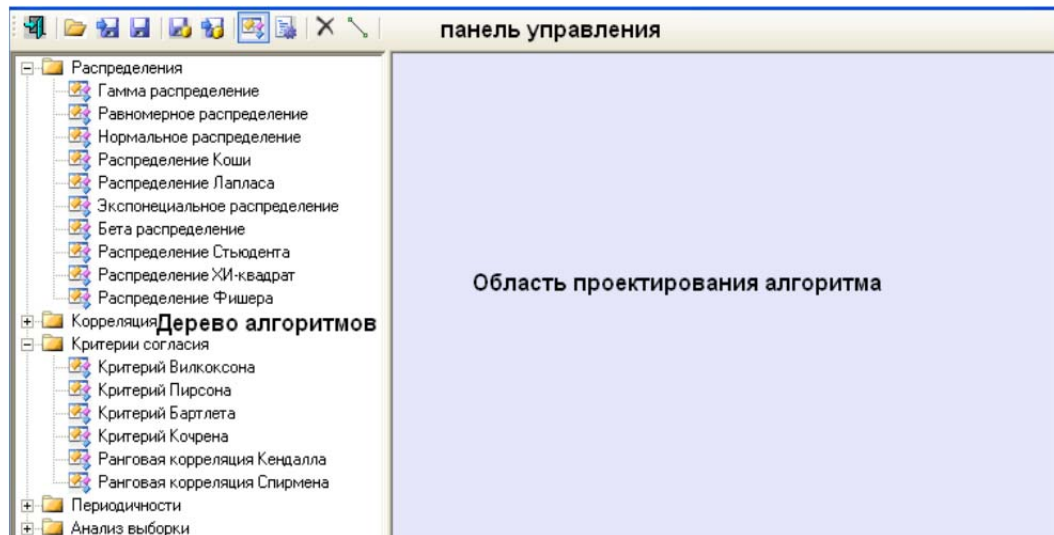
Созданный алгоритм может сохраняться как в базу данных так и в файл в XML формате.

Физически алгоритм представляет собой логическое описание связей между базовыми алгоритмами где выходные значения одних алгоритмов являются входными других. В результате получается цепочка обработки статистических данных. Также при создании алгоритма могут указываться значения входных данных по умолчанию (например уровень значимости при проверке гипотез) а также выходные значения необходимые для вывода пользователю в качестве результата функционирования.

Для решения статистических задач в условиях малых выборок или значительной неопределенности на основе регуляризирующего байесовского подхода (РБП) [1,2] требуется система, которая позволила бы получать входные данные из различных источников, обрабатывать их и выводить в заданном формате. Такой системой является «Байесовская математическая статистика». Система предназначена для дополнительной обработки данных полученных с платформы «Инфоаналитик» (свидетельство об официальной регистрации программы для ЭВМ № 2004611956 от 25.09.2004).

Основная идея системы заключается в том, чтобы объединить классический статистический подход с РБП. Для того чтобы этого добиться в рамках системы вводятся два типа данных: обычные дей-

ствительные числа и регуляризирующие байесовские оценки (РБО). Это позволяет применять статистические алгоритмы, как на четких, так и на нечетких данных.



Система устроена по аналогии с таким известным пакетом как Statistica, но в отличие от него «Байесовская математическая статистика» позволяет решать задачи в условиях значительной неопределенности. Это достигается встроенной поддержкой регуляризирующих байесовских оценок.

Также имеется библиотека алгоритмов, организованная в виде дерева. Все алгоритмы настраиваются единообразно, путем задания всех входных и выходных параметров в стандартном диалоге. Настроенный алгоритм является уже не простым действием по обработке данных, которое активизирует пользователь, а некоторым алгоритмическим объектом, который содержит в себе всю информацию необходимую для обработки и может запускаться в автономном режиме, например, при поступлении новых данных в режиме реального времени

Для обработки данных необходимо выбрать алгоритм из библиотеки и настроить его в соответствии с теми данными, которые необходимо обработать, после чего его можно запустить и увидеть результат его работы.

Алгоритмы подключаются к системе как независимые компоненты, что позволяет разрабатывать их отдельно от системы.

В системе имеется возможность визуализации данных и результатов их обработки, что позволяет их представить наглядно в виде диаграмм и графиков, когнитивных образов.

Все вышеперечисленные объекты (рабочие листы, алгоритмические объекты, диаграммы) можно сохранить в базу данных, организованную в виде дерева. Это избавляет от повторных действий по настройке алгоритмов и диаграмм. Кроме того, в базе данных хранятся связи между этими объектами, что позволяет организовать процесс обработки данных в единую систему.

На базе системы «Байесовская математическая статистика» планируется разработать функциональность позволяющую организовывать вызовы алгоритмов в последовательные цепочки, чтобы автоматизировать процесс обработки данных и метрологического сопровождения.

В совокупности с платформой «Инфоаналитик» система «Байесовская математическая статистика» удовлетворяет основным требованиям предъявляемых к современным информационным технологиям интеллектуальных измерений:

1. Обеспечение контроля качества получаемых результатов и определение их основных метрологических показателей в виде значений точности, надежности и достоверности, составляющих принципиальную основу метрологического сопровождения.

2. Учет специфики объекта измерения, вида и степени определенности априорной и получаемой экспериментальной информации о нем при выборе методов и средств измерений.

3. Возможность оптимизации планирования проведения измерительного эксперимента в соответствии с метрологическими требованиями и ограничениями.

4. Обеспечение полноты и достаточности получаемых решений на основе учета качественной и количественной информации.

5. Способность динамической адаптации технологии к меняющимся условиям и свойствам объекта измерений в процессе эксперимента.

6. Возможность самообучения и на основе этого самоорганизации информационно-измерительных систем.

7. Создание развивающихся и саморазвивающихся информационных технологий и информационно-измерительных систем.

1.4. Мягкие технологии аудита и управления сложными системами в условиях неопределенности

Под «мягкими» технологиями понимаются информационные технологии, обладающие свойствами изменения своей структуры

в зависимости от меняющихся внешних условий и состоящие из совокупности альтернативных технологий с соответствующими им вероятностями их реализации.

Процессы оценки состояния и управления распределенными технологическими системами, например, топливно-энергетическими комплексами, транспортными магистралями или территориями являются типичными примерами контроля и управления сложными распределенными объектами с меняющимися во времени и в пространстве свойствами. При активном взаимодействии таких объектов с окружающей средой в условиях значительной неопределенности информации, ситуаций, меняющихся ограничений (например, по нормативно-методической или правовой, баз), критериев и требований возникает задача реконфигурации моделей, структуры и функций систем мониторинга, аудирования, управления объектами в режиме функционирования самих объектов. При эксплуатации указанных информационных систем это обеспечивает постоянное поддержание адекватности используемых в них моделей объектам с целью обеспечения эффективности получаемых решений. Например, для управления предприятиями топливно-энергетического комплекса по неполной и неточной информации, собираемой системой мониторинга, необходимо оценить техническое состояние объекта, экологичность производства и его влияние на окружающую среду, производственные ресурсы, кадровую, финансовую и рыночную ситуации и другие компоненты деятельности предприятия и окружающей среды. Очевидно, что для выработки эффективной стратегии управления таким объектом необходимо иметь в составе системных средств средства для реконфигурации модели объекта с достаточной степенью достоверности и контролируемостью риска принятия решений в темпе поступления новой информации об объекте и среде. Таким образом, учитывая вышеперечисленные требования и ориентируясь на свойства интеграции, методологичности и саморазвития методологической основы регуляризирующего байесовского подхода (РБП) и информационных технологий на его основе (байесовских интеллектуальных технологий – БИТ) представляется целесообразным использовать их для создания развивающихся систем мониторинга, аудита и управления.

Согласно современным теоретическим работам философии настоящее время характеризуется новым концептуальным представлением принципов управления сложными динамическими системами.

«Под управлением понимается не жесткая детерминированная схема, а «мягкие формы управления». Фактически доминирующими видами управления становятся разнообразные виды управления через среду» [1] на основе принципов полисубъектного управления по типу «субъект-полисубъектная среда».

Исходные положения и рефлексивные модели полисубъектного управления были впервые сформулированы В.А. Лефевром [2]. В настоящее время на этих принципах активно развивается концепция постнеклассической рациональности и управления на основе кибернетики третьего порядка «человекообразными саморазвивающимися средами».

Определение понятия «мягкого управления» и формальные модели на основе схем саморазвития были даны в работах автора [3–5]. На основании этих формальных моделей были разработаны информационные технологии и алгоритмы «мягкого управления» сложными динамическими системами техногенного и социально-экономического типов.

Процессы оценки состояния и управления распределенными техногенными системами, например, топливно-энергетическими комплексами, транспортными магистралями или территориями являются типичными примерами контроля и управления сложными динамическими системами с меняющимися во времени и в пространстве свойствами. При активном взаимодействии таких объектов с окружающей средой в условиях значительной неопределенности информации, ситуаций, меняющихся ограничений (например, по нормативно-методической или правовой, баз), критериев и требований возникает задача реконфигурации моделей, структуры и функций систем мониторинга, аудирования, управления объектами в режиме функционирования самих объектов. При эксплуатации указанных информационных систем это обеспечивает постоянное поддержание адекватности используемых в них моделей объектам с целью обеспечения эффективности получаемых решений.

Управленческая практика выдвигает новые задачи полисубъектного управления. Так, в процессе управления предприятиями по неполной и неточной информации, собираемой системой мониторинга, необходимо одновременно оценить и скорректировать в заданном направлении техническое состояние управляемых систем, экологичность производства и его влияние на окружающую среду, производственную политику, кадровую, финансовую и рыночную ситуации и другие процессы деятельности предприятия и окружающей среды.

Очевидно, что для выработки эффективной стратегии управления таким объектом необходимо иметь в составе системных средств средства для реконфигурации модели объекта с достаточной степенью достоверности и контролируемостью риска принятия решений в темпе поступления новой информации об объекте и среде.

Учитывая вышеперечисленные требования и ориентируясь на свойства интеграции, метрологичности и саморазвития методологической основы регуляризирующего байесовского подхода (РБП) и информационных технологий на его основе (байесовских интеллектуальных технологий – БИТ) представляется целесообразным использовать их для создания развивающихся систем мониторинга, аудита и управления по схеме «мягкого управления».

Очевидно, все эти задачи могут быть разделены на три основные группы:

- оценивание свойств и характеристик объекта, экосистемы и ее компонент, а также моделирование их эволюции, восстановление ретроспективы развития и прогнозирование состояний и ситуаций;
- контроль (аудирование) и нормирование состояния техногенного объекта экосистем или их характеристик; к этому этапу относятся все задачи различных аудитов (качества продукции и производства, экологичности производств, персонала, энергетических показателей и др.);
- генерация оптимальных управленческих решений и управляющих рекомендаций.

В данном докладе рассмотрены возможности создания таких саморазвивающихся систем на основе моделей и шкал с динамическими ограничениями (ШДО) [1].

Системы на основе БИТ имеют поля решений, деревьев факторов и их свойств в виде динамических компактов ШДО для объектов S_t^O , окружающей среды S_t^E , или объекта, взаимодействующего со средой S_t^{OE} . При оценивании свойств или ситуаций на основе методологии БИТ концептуальная запись ШДО имеет вид:

$$S_t^{OE} = S_t^O \overset{z}{*} S_t^E \overset{z}{*} S_t^{O_L} \overset{z}{*} S_t^{O_K} \overset{z}{*} S_t^{O_U} \quad (51)$$

где $S_t^{O_L}$ – ШДО ограничений в пространстве и во времени;

$t = \overline{1, T}$ - временной интервал, определяющий динамику ШДО;

$S_t^{O_k}$ – ШДО критериальной базы;
 $S_t^{O_v}$ – ШДО требования (в том числе метрологические требования) и условия постановки задачи;
 $\overset{z}{*}$ – символ свертки с логикой z (при $z = 1$ и параметрическом семействе логик свертка соответствует свертке с вероятностной логикой).

Для синтеза ШДО разработана методологическая база и информационная метатехнология, которые позволяют достаточно просто осуществлять переход к новому пространству существования решений, оценок или моделей. Для интеграции такого пространства в существующие рабочие поля и обеспечения квазиустойчивости решений в условиях неопределенности новой информации производится метризация пространства на основе ШДО. В этой методологии могут меняться требования, ограничения, условия решения задачи, критериальная база, логика вывода и принятия решений, априорная информация, информационно-технологическая и техническая базы. Поэтому выражение (52) по сути представляет краткую формализованную запись концепции реконфигурации моделей сложного объекта и среды для условий неопределенности.

В силу возможности управления не только данными, но и знаниями [Сборники докладов конференции SCM-98, SCM-99, SCM-2000, SCM-2001] в системах на основе РБП, а также открытости их для включения в информационные базы новых данных моделей, алгоритмов и рекомендаций на основе шкал с динамическими ограничениями теоретических и практических знаний обеспечиваются требования реконфигурации моделей (52). Каждое пополнение такого рода сопровождается в системе процессом метрологического обоснования вновь принятой информации. Для отражения параметрической, функциональной или системной информации строится иерархическая структура гипер-ШДО HS в виде композиции параметрических $S_t^{O_p}$, функциональных $S_t^{O_f}$ и системных (ситуационных) $S_t^{O_s}$ ШДО:

$$HS_t^O = S_t^{O_p} \overset{z}{*} S_t^{O_f} \overset{z}{*} S_t^{O_s} \quad (53)$$

Концептуальная запись, для синтеза всех типов ШДО основным принципиально важным аспектом метрологического обоснования является метрология знаний, в том числе и знаний, представленных в

виде аналитических зависимостей, алгоритмов, лингвистических выводов, описания ситуаций и сценариев их развития, что реализуется на базе ШДО [Сборники докладов конференции SCM-98, SCM-99, SCM-2000, SCM-2001]. Применение ШДО позволяет контролировать качество и метрологические характеристики (МХ) (точность, надежность, достоверность), энтропию и риск получаемых решений, а также синтезировать технологию с требуемым качеством.

Разработаны ШДО для оценки числовых S_i^N и лингвистических S_i^{Lg} переменных, динамических переменных $S_i^S(t=\overline{1,T})$ и многомерных процессов $HS_i^S(t=\overline{1,T})$ ситуаций $S_i^S(t=t_i)$, сценариев развития ситуаций $S_i^S(t=\overline{1,T})$, рекомендаций $S_i^D(t=f.)$, стратегий управления $HS_i^D(t=\overline{1,T})$ ШДО взаимосвязи объектов и ситуаций S_i^{SS} . Последний тип шкал позволяет создавать гипершкалы для интегральных характеристик (индексов, индикаторов), восстанавливать их или прогнозировать.

Для проведения аудита состояний или ситуаций дополнительно к перечисленным формируются эталонные шкалы указанных типов S_i^C или гипершкалы HS.

Эталонная ШДО для аудита S_i^C представляет собой регуляризованную информацию о нормах, стандартах и степенях соответствия установленным пределам показателей характеристик объекта. В частности, проведение проверки на соответствие стандартам ISO 9000 или ISO 1400 реализовано на методологической и информационно-технологической базе байесовских интеллектуальных технологий средствами ШДО. Это необходимое каждому предприятию программное средство для проведения самоаудита, обеспечивающее значительную экономию финансовых и временных затрат. Концептуальное уравнение такого **интеллектуального аудита**, определяемого в данной работе как аудита, основанного на производстве, использовании знаний и управлении знаниями в условиях неопределенности (при использовании БИТ – **байесовского интеллектуального аудита**) может быть записано в оптимизационной форме следующего вида:

$$\left\{ h_i^c \mid \{MX\}_i^c \right\} = \left\{ \arg \text{extrC } S_i^{OE} \overset{z}{*} S_i^{CL} \overset{z}{*} S_i^{CU} \overset{z}{*} S_i^{CK} \mid (X_i; X_i^A; \Phi_i^A; \Phi_i; \Phi_i^{CA}; \Phi_i^C) \right\} \quad (54)$$

где $X_i; X_i^A$ – массивы текущей и априорной информации;

$\Phi_t^A; \Phi_t$ – информационная технологическая база в текущий и предыдущий моменты времени;

$\Phi_t^{CA}; \Phi_t^C$ – расчетно-методическая основа аудита, представленная в виде набора информационных технологий контролирующих (аудиторских) методик, которые интегрируются средствами байесовских интегрирующих технологий;

S_t^{CL} – ШДО ограничений при проведении аудита;

S_t^{CU} – ШДО условий аудита;

S_t^{CK} – ШДО критериальной базы аудита;

C – байесовское решающее правило с гибкой логикой.

Если применяется отличная от вероятностной логика байесовского правила, то такое правило называется **мягким байесовским правилом (Soft Bayesian Rule)**. Аудит, основанный на применении ШДО с мягким байесовским правилом можно определить как «**мягкий**» (*soft audit*) *аудит*.

Основные отношения, реализованные на ШДО, позволяют интегрировать разнотипные информационные потоки на основе свертки метрических пространств ШДО. Изменение компакта и структуры носителей ШДО в зависимости от меняющихся внешних условий, требований, ограничений, целей или критериев решаемых задач позволяет в режиме реального времени изменить структуру гипершкалы объектов или ситуаций, а вместе с ней и модель объекта аудита или управления. Таким образом, каждому моменту времени соответствует своя модель и гипершкала, что обеспечивает их адекватность развивающемуся объекту и меняющейся внешней среде. Для каждого момента времени в технологиях БИТ существуют априорные шкалы и модели, текущие и апостериорные как результат свертки двух предыдущих типов шкал.

Для обеспечения процесса интеграции информационных ресурсов и технологий на ШДО необходима метризация пространства решений или компакта гипершкалы в соответствии с требованиями и ограничениями задачи.

Для этого в технологиях БИТ предусмотрена параллельная ветвь метрологического синтеза ШДО. Свойство инертности экосистем в определенных пределах внешних воздействий и шкалирование ситуаций по РБП позволяют перейти от жесткого управления по оценкам отдельных контролируемых параметров этих экосистем к более глу-

бокому комплексному управлению на основе оценки ситуаций. **Интеллектуальное («мягкое» при гибкой логике) ситуационное или стратегическое управление** на основе методологии РБП и БИТ дает возможность предусмотреть влияние результатов предпринимаемого регулирования на все компоненты объекта и экосистемы и предупредить появление кризисных или чрезвычайных ситуаций, а при необходимости своевременно скорректировать стратегию управления, что является основой обеспечения устойчивого функционирования или развития объекта. Однако для обеспечения устойчивости управления в условиях неопределенности (неточности, неполноты и нечеткости) информации необходимо достоверно оценить ситуацию.

Предлагаемый в данном проекте регуляризирующий байесовский подход (РБП) и методология байесовских интеллектуальных технологий (БИТ) дают уникальную возможность обеспечить получение таких оценок и требуемого качества их определения. На основе этого подхода производится квантификация пространства ситуаций, рекомендаций или управляющих воздействий, что позволяет сформировать их репрезентативное множество в виде многомерных и интегральных шкал ситуаций, реперами которых являются градации признаков или типовых классов ситуаций, отстоящие друг от друга на заданное расстояние.

Процесс принятия решений может быть реализован как процесс преобразования таких взаимосвязанных шкал на основе получения объективных и достоверных: знаний о контролируемом объекте и окружающей его среде, знаний об их текущем состоянии, динамики развития ситуаций, тенденциях в эволюции объекта и среды. Процесс получения таких знаний и управление на его основе, очевидно, можно назвать **интеллектуальным управлением на основе БРП/БИИ/БИТ**. Если логика такого управления реализуется на основе мягкого байесовского правила со сменной логикой получения решений, то такое управление можно определить как **«мягкое» управление**.

Гипершкала для проведения интеллектуального управления может быть представлена в следующем виде:

$$HS_t^D = HS_t^C \overset{z}{*} HS_t^O \overset{z}{*} S_t^D \overset{z}{*} S_t^{DL} \overset{z}{*} S_t^{DK} \overset{z}{*} S_t^{DU} \quad (55)$$

где для процесса управления: S_t^D – ШДО рекомендаций, управляющих воздействий стратегии управления;

S_t^{DL} – ШДО ограничений в пространстве и во времени;
 S_t^{DK} – ШДО критериальной базы управления;
 S_t^{DU} – ШДО требования (в том числе метрологические требования) и условия постановки задачи управления.

Концептуальное уравнение байесовского интеллектуального и мягкого управления приведено здесь в форме:

$$\{h_t^c | \{MX\}_t^c\} = \{\arg \text{extr} C(HS_t^D) | (X_t; X_t^A; \Phi_t^A; \Phi_t; \Phi_t^{DA}; \Phi_t^D)\} \quad (56)$$

где $\{h_t^c | \{MX\}_t^c\}$ – гибкая стратегия байесовского интеллектуального управления с комплексом метрологических характеристик;

$\Phi_t^{DA}; \Phi_t^D$ – расчетно-методическая основа управления, представленная в виде набора информационных технологий контрольных (аудиторских) методик, которые интегрируются средствами байесовских интегрирующих технологий.

Спецификой такого управления является отсутствие соблюдения в полной мере принципов наблюдаемости, идентифицируемости и, как следствие, полной управляемости объектом. В силу непознаваемости природных и техногенных объектов, помещенных в природную среду в этой ситуации действуют принципы квазинаблюдаемости (неполной, нечеткой, неточной наблюдаемости), квазиидентифицируемости и, вообще говоря, квазиуправляемости.

Основной целью интеллектуального управления на основе РБП в условиях значительной неопределенности является получение гибких, нечетких, **«мягких» стратегий управления** на основе знаний о регулируемых и регулирующих факторах, оценка их свойств, состояний и взаимовлияние сложных объектов искусственной и природной среды, определение динамики и тенденций развития объектов моделирования и в условиях меняющегося влияния среды их функционирования, а также выработка рекомендаций, обеспечивающих устойчивое функционирование, результаты чего имеют полное метрологическое обоснование их качества. Такие технологии управления позволяют переходить с одной стратегической линии управления объектом или территорией, осуществлять стратегическое планирование развития объекта (предприятия или территории) в зависимости от новых условий,

формируя в каждый момент наиболее эффективное и полное управленческое решение.

Многие современные методы управления сложными системами характеризуются мягкими формами управления, которые создаются на основе мягких моделей, информационных технологий мягких измерений, мягкого контроля и аудита [1–3]. Такие методологические решения, связаны с самой природой сложных систем, определяющей высокую степень неопределенности информации об их свойствах, с одной стороны и, с другой стороны, активным взаимодействием этих систем с меняющейся окружающей средой, существенно влияющей на них. Все это в полной мере относится к управлению полисистемами, составляющими значительную часть современной экономической среды [2, 3].

В настоящее время, характеризуемое кризисными ситуациями в экономике, управление производственными и социально-экономическими системами основывается на риск-ориентированных подходах. Риск-ориентированные методы управления призваны обеспечивать устойчивость функционирования управляемых систем в нестабильных средах. Это, в первую очередь, относится к полисистемам, состоящим из ряда автономных систем, при слиянии и взаимодействии которых риски усиливаются и умножаются. При этом основными задачами управления являются задачи идентификации, измерения и аудиторной оценки рисков. Это является по сути концепцией «оборонительной» экономической политики. Однако, непрерывное и эффективное развитие таких систем связано с созданием и реализацией потенциалов систем. Например, при создании инвестиционного проекта и разработке его бизнес-плана особое внимание уделяется оценке финансовых поступлений при реализации проекта, обеспечивающих эффективность и окупаемость проекта. При отсутствии или недостатке реальных договоров на поставку новой продукции или услуг проекта, обеспечивающих финансовые потоки для окупаемости проекта, существует практика «выращивания потребителей», которая состоит в дополнительных детальном маркетинговых исследованиях и реализации практических действий по созданию рынков и потребителей продукции инвестиционного проекта. Таким образом создаются дополнительные потенциалы проекта. Или, например, специалист реализует профессиональную самоподготовку, проходит дополнительное обучение с целью создания потенциала выполнения других видов

работ, продвижения по основной деятельности. Предприятие дифференцирует номенклатуру продукции с целью усиления производственного потенциала и занятия новых рыночных ниш. Такая экономическая политика, основанная на усилении и создании экономических потенциалов, является «наступательной». Соединение в единую политику указанных двух создаст условия для гибкого и эффективного управления сложными системами при решении разнообразных задач в различных ситуациях.

В связи с этим целесообразно разработать концепцию комплексного управления сложными системами, в частности полисистемами, на основе интеграции риск-ориентированного и потенциал-ориентированного подходов. Определим этот подход к управлению как потенциал-риск-ориентированный менеджмент (ПРОМ/PROM).

Концепция PROM должна предусматривать идентификацию, измерение, контроль потенциалов и рисков систем в режиме мониторинга, а также их прогнозирование на каждом этапе реализации функций систем. В связи со значительной степенью неопределенности и нестабильности реальных систем, а также неустойчивостью и изменчивостью связей между ними и факторами окружающей среды, при создании концепции PROM необходимо использовать в качестве базовых методологий, методологии ориентированные на указанные специфические свойства систем.

К числу таких подходов принадлежит регуляризирующий подход (РБП). Рассмотрим информационные технологии решения указанных выше задач управления в рамках концепции PROM на основе байесовских интеллектуальных технологий (БИТ).

В число основных информационных мягких технологий PROM на основе БИТ входят:

1. Технологии идентификации и измерения потенциалов и рисков систем на основе байесовских интеллектуальных измерений (БИИ), в частности, байесовских мягких измерений (БМИ).
2. Технологии нормирования и контроля (аудита) потенциалов и рисков на основе БИТ.
3. Технологии оценки и интерпретации динамики потенциалов и рисков свойств системы и среды на основе БИТ и регуляризированных байесовских оценок (РБО).
4. Технологии оценки тенденций и трендов потенциалов и рисков, а также свойств систем на основе БИТ.

5. Технологии моделирования потенциалов и рисков в виде нечетких (мягких) моделей на основе БИТ.

6. Технологии прогнозирования потенциалов и рисков на основе БИТ и их регуляризированных байесовских оценок (РБО).

7. Технологии генерации управленческих решений для управления потенциалами и рисками основе БИТ.

8. Технологии генерации управленческих решений PROM для управления функциями автономных систем на основе БИТ.

9. Технологии генерации управленческих решений PROM для управления интегральными потенциалами и рисками полисистемы на основе БИТ.

10. Технологии генерации управленческих решений для управления интегральными функциями полисистем основе PROM.

11. Технологии генерации управленческих решений для управления влиянием внешней среды на основе БИТ.

12. Технологии генерации управленческих решений для управления созданием целесообразных потенциалов на основе БИТ.

13. Технологии генерации управленческих решений для управления реализацией созданных потенциалов на основе БИТ.

14. Технологии развития субъектов PROM.

15. Технологии реконструкции и саморазвития систем PROM на основе БИТ.

Этот перечень технологий, реализуемый в указанной последовательности представляет собой мета-технологию PROM.

Для идентификации рисков и потенциалов сложных систем необходимо уточнить их определения как объектов идентификации и измерения. Так, риск R безотносительно его прикладной направленности в известных информационных источниках определяется как вероятность наступления нежелательного события. Очевидно, что в этом определении вероятность носит субъективный, а не частотный характер. По аналогии с определением риска определим *потенциал P* как *вероятность наступления благоприятного события, где вероятность имеет также субъективный характер.*

Этап идентификации состоит в разработке идентификационных моделей потенциалов и рисков сложных систем, которые затем должны быть использованы при измерении и управлении потенциалами и рисками. Для целенаправленного управления потенциалами и рисками сложных систем необходимо создать вначале модели самих си-

стем, а затем модели потенциалов и рисков, соответствующих им. Изменение свойств сложных систем и взаимосвязей с внешней средой задает необходимость постоянной коррекции модельных представлений об объекте, оценок его свойств, критериев и других компонентов информационного пространства моделей. Иными словами, модель должна обладать свойством адаптации не только к конкретному объекту исследования, но и к конкретной возникшей ситуации в деятельности данного объекта. Таким свойством обладают модели и методики, созданные в работе на основе БИТ, названные в терминологии БИТ моделями с динамическими ограничениями (МДО). По этим технологиям разработанная модель представляется в виде дерева факторов потенциалов и рисков, соответствующих цели управления, как показано на **рисунке 24**.

В рамках идентификации и в зависимости от природы происхождения потенциалы и риски (П&R) можно разделить на три основных типа:

1. П&R *системности*, возникающие при слиянии систем и обусловленные свойством эмерджентности интегрирующей системы.
2. П&R *нестабильности*, связанные с изменчивостью факторов и ситуаций самой системы и внешней окружающей среды.
3. П&R *неопределенности*, связанные с информационной неполнотой, неточностью, нечеткостью, недостоверностью и противоречивостью данных и знаний о системах и влияющих факторах внешней среды.

В зависимости от типа потенциалы и риски регулируются специальными управляющими воздействиями, ориентированными на специфику типов.

Так П&R системности регулируются введением дополнительных факторов в идентификационную модель. П&R нестабильности контролируются путем введения динамических идентификационных моделей. П&R неопределенности оцениваются и управляются путем применения методологии РБП и БИТ, ориентированных на эти условия.

В условиях ситуационной нестабильности и информационной неопределенности можно представить концептуальную модель П&R каждой автономной системы в любой момент времени t в виде динамической модели $G^{(авт)}_{it}$:

$$G^{(авт)}_{it} = G^{(П)}_{it} * G^{(R)}_{it} * G^{(E(П\&R))}_{it} | Y_{it}, \quad (57)$$

где $G^{(П)}_{it}$ – модель целевых потенциалов i -той автономной системы;

$G_{it}^{(R)}$ – модель рисков i -той автономной системы;
 $G_{it}^{(E(\Pi \& R))}$ – модель потенциалов и рисков, связанных с влиянием окружающей среды;

Y_{it} – модель условий моделирования, представляющих собой байесовскую свертку априорной информации, требований и критериев, и ограничений;

* – символ свертки свойств и показателей составляющих моделей.

Концептуальную модель PROM для полисистемы $G_t^{(c)}$ можно записать в аналогичном виде:

$$G_t^{(c)} = * (G_{it}^{(авт)}) * G_t^{(сис)} * G_t^{(E(\Pi \& R))} | Y_t * G_t^{(c)} * G_t^{(c)}, \quad (57)$$

где $G_{it}^{(\Pi)}$ – модель целевых потенциалов полисистемы;

$G_t^{(R)}$ – модель рисков полисистемы;

$G_{it}^{(E(\Pi \& R))}$ – модель потенциалов и рисков $G_t^{(c)}$, связанных с влиянием окружающей среды;

Y_{it} – модель условий моделирования.

Для измерения потенциалов и рисков по неполной неточной и недостоверной количественной и качественной информации, а также для повышения устойчивости получаемых оценок разработаны методики измерения показателей потенциалов и рисков на основе технологий байесовских интеллектуальных измерений (БИИ) на шкалах с динамическими ограничениями (ШДО).

В методологии БИИ и программном комплексе «Инфоаналитик» на ее основе созданы специальные прикладные подсистемы, последовательное использование которых позволяет выполнять все этапы создания концептуальной модели и получения оценки потенциалов и рисков на основе БИТ. Подсистемы позволяют построение модели объекта и среды с динамическими ограничениями (МДО); создание ШДО для определения свойств объектов и среды в условиях неопределенности информации для множества альтернативных решений интегрального показателя с комплексами метрологических характеристик. При этом интегральная ШДО сохраняет вид стандартной шкалы с таким же набором реперов; построение динамических моделей БИТ и другие частные подзадачи.

Вид ШДО для измерения потенциалов показан на **рисунке 24**.

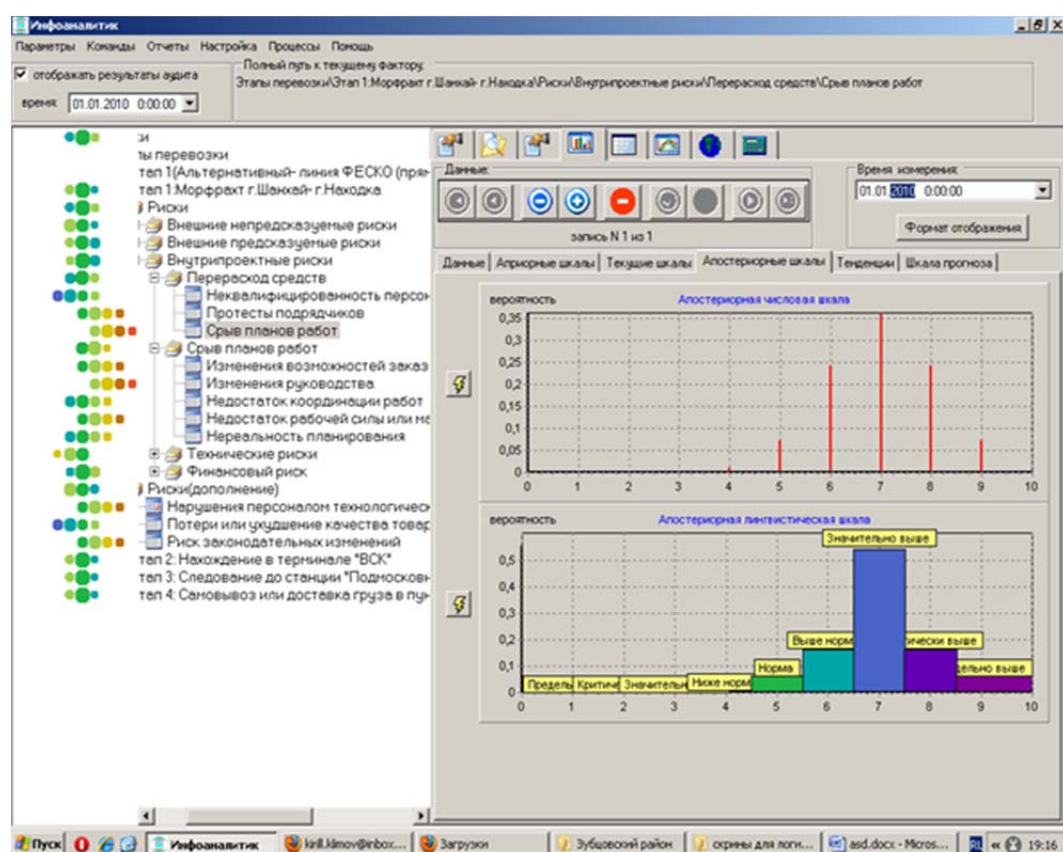


Рис. 24. Идентификационная модель и ШДО для измерения потенциалов

ШДО для измерения потенциала имеет носитель шкалы, диапазон которого расположен в интервале от 1 до 9. Значимым считается потенциал, значение которого выше 5. Риски и потенциалы для каждого j -ого фактора связаны между собой соотношением $P_j = 9 - R_j$.

Результат измерения потенциалов по методологии БИИ представляется в нечеткой форме в виде регуляризированной байесовской оценки (РБО) $\{h_k\}$, которая по сути представляет собой оценку условного закона распределения $f_{jt}(h)|G^{(c)}_t$ потенциала для j -ого фактора модели $G^{(c)}_t$ в виде столбчатой диаграммы. Разработаны методы и средства, для аппроксимации этих оценок законов распределения моделями типовых унимодальных кривых плоскости Пирсона. При формирование интегральных оценок потенциалов $f(h)$ и рисков осуществляется байесовская свертка полученных моделей законов распределения и закон распределения интегральной оценки потенциалов полисистемы может быть записан в виде:

$$f_t(h)|G^{(c)}_t = * f_{jt}(h)|G^{(c)}_t \quad (58)$$

На **рисунке 24** показаны также результаты аудита потенциалов и рисков в виде когнитивной графики, иллюстрирующей оценки в виде разноцветных кружков, цвет которых соответствует значению оценки. Потенциалы, значения которых достигают высоких значений (выше 7) имеют синюю или фиолетовую окраску кружков.

В качестве оценки может быть взята средневзвешенная по вероятности оценка потенциала. Качество получаемых оценок определяется комплексами метрологических характеристик, которых показатели точности, надежности и достоверности решений определяются в виде лингвистических оценок на основе условных фидуциальных вероятностей $P(h_j)|f_t(h)|G^{(c)}_t$. **Рисунок 25** иллюстрирует динамические оценки потенциалов

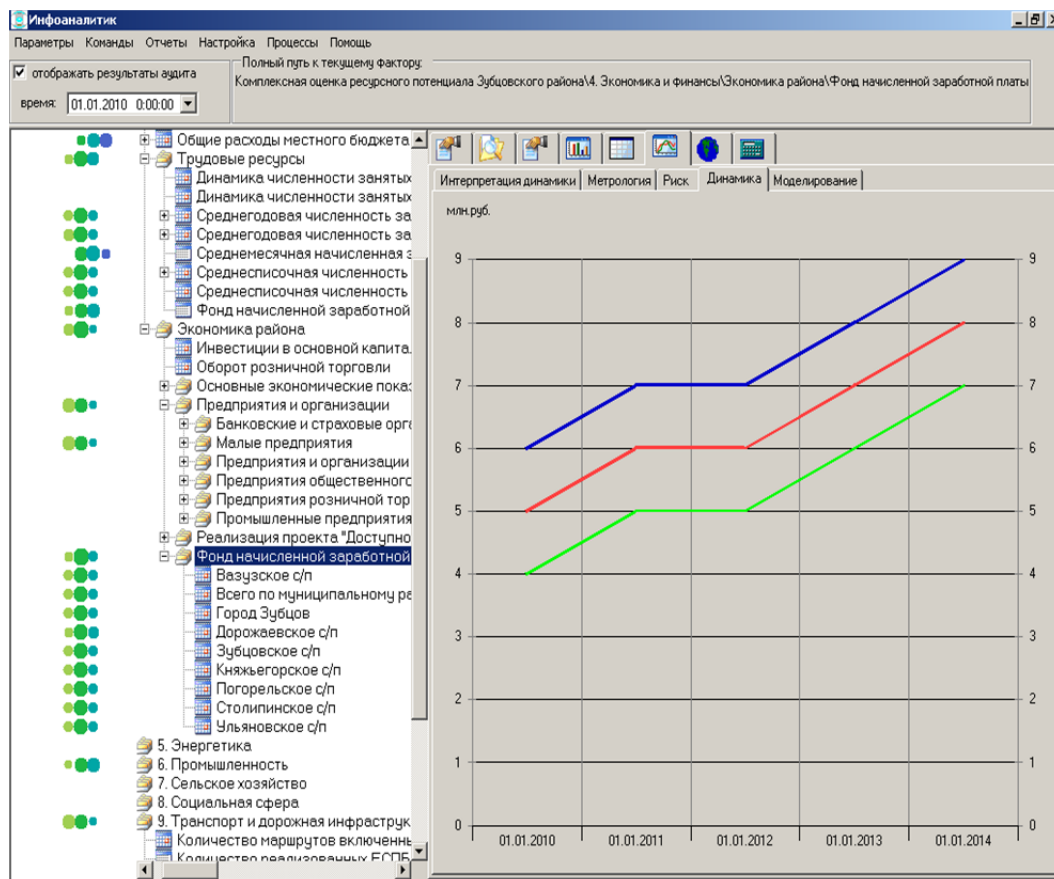


Рис. 25. Нечеткая динамическая модель потенциала

Показатели риска определяются аналогичным образом, но ШДО для измерения рисков является инвертированной по отношению к ШДО измерения потенциалов.

Методика управления потенциалами и рисками реализует все этапы концепции.

1.5. Решение прикладных задач на основе регуляризирующего байесовского подхода

1.5.1. Опыт применения РБП и БИТ для решения измерительных и управленческих задач

Методология, информационные технологии и системы на основе РБП нашли применение в различных сферах практической деятельности. Приведенный ниже список некоторых научных работ и проектов, выполненных на основе РБП-технологий дан не столько для того, чтобы убедить читателя в эффективности таких решений, но для того, чтобы читатель мог определить интересующие его сферы прикладных задач и детально затем ознакомиться с полученными решениями.

- Информационно-аналитическая система для стратегического планирования экономического и социального развития государств регионов и городов Черноморского региона.

- Разработка инвестиционной программы для развития теплоэнергетики г. Краснодара.

- Информационно-аналитическая система Электронное правительство – Информационные системы: «Страна» – «Регион» – «Губернатор» – «Мэр» – «Директор», адаптированные к политико-экономическим особенностями территориальной специфике региона.

- Распределенная информационно-аналитическая система для управления топливно-энергетическим комплексом региона для ТЭК Томской области.

- Сеть распределенных информационных центров для регулирования энергоснабжения и использования водных ресурсов стран Центральной Азии (проект Экономической комиссии ООН «СПЕКА» для развития экономики стран Центральной Азии).

- Информационно-аналитическая система для мониторинга технологий в рамках программы ЕК ООН «Чистый уголь».

- Обучение специалистов Министерств стран Центральной Азии и специалистов энерго- и водоснабжения стран Центральной Азии (проект ЕК ООН).

- Распределенная геоинформационная аналитическая сеть для управления биотопливными ресурсами и генерацией энергии из воз-

обновляемых источников республики Беларусь (проект ЕК ООН, Комитет по устойчивой энергетике).

- Обучение специалистов по программе «Энергоэффективность-21» (проект ЕК ООН).

- Информационно-аналитическая система интерпретации и прогнозирования производственных и эколого-экономических ситуаций в регионе, сценариев их развития (предприятие «Адмиралтейские верфи»).

- Информационно-аналитическая система для мониторинга территории Калининградской области.

- Геоинформационная система с картированием эколого-экономических ситуаций в Балтийском регионе.

- Информационно-аналитическая система для управления природными ресурсами Байкальского региона. Кадастры природных ресурсов Байкальской природной территории.

- Информационно-аналитическая система для оценки эколого-экономической ситуации в регионе. Передвижная экологическая станция г. Макеевка, Донецкой области.

- Информационно-аналитическая система для мониторинга состояния Черного моря.

- Информационно-аналитическая система для сохранения биологического разнообразия стран Черноморского региона.

- Информационно-аналитическая система для предупреждения чрезвычайных ситуаций в Ямало-Ненецком округе.

- Информационно-аналитическая система оценки воздействия на окружающую среду. На Байкальской природной территории.

- Информационно-аналитическая система расчета загрязнений воздушной и водной среды, почв, растительного покрова Байкальского региона.

- Электронная красная книга, летопись природы, электронная библиотека памятников природы и особо охраняемых природных территорий. Пинежского заповедника.

- Информационно-аналитическая система энергоаудита и энергоменеджмента предприятий, жилищно-коммунального хозяйства, территорий г. Краснодара.

- Информационно-аналитическая система для управления предприятиями и организации предпринимательской активности. Интеллектуальное рабочее место менеджера. Электронный документооборот. Электронная библиотека специалиста. Управление ресурсами и электронная логистика. Для предприятия ОАО «Лентрансгаз».

- Информационно-аналитическая система для управления инвестиционными проектами в Черноморском регионе.
- Информационно-аналитическая система для управления персоналом организаций. Электронный документооборот по учету труда. Мониторинг и аудит деятельности сотрудников компании. Контроль исполнения и управление заданиями. Дистанционное обучение. Интеллектуальное рабочее место специалиста. Управление квалификацией сотрудников. Кадровый аудит и консалтинг (Центрохотконтроль).
- Информационно-аналитическая система для управления имуществом и инвентаризации (ОАО Роснефтегазстрой).
- Информационно-аналитическая система для оптимизации хозяйственной деятельности РФ (Минсельхоз, Центрохотконтроль).
- Информационно-аналитическая система для управления бизнесом. Электронная коммерция. Интеллектуальная система управления бизнес-рисками. Управление отношениями с клиентами. Генерация оптимальных бизнес-технологий. Эффективные маркетинговые коммуникации.
- Создание корпоративных интеллектуальных сетей. Управление корпоративными знаниями. Управление информационными потоками.
- Информационно-аналитическая система экологического аудита предприятий и менеджмента согласно ISO 14001, ISO14001P, EMAS1893/36.
- Интеллектуальные системных управления эффективным использованием энергетических и водных ресурсов для Стран Центральной Азии (ПРООН. ЕК ООН).
- Проект ООН по созданию региональной информационной сети (Казахстан, Киргизия, Таджикистан) для национальных институтов, решающих задачи энерго- и водосбережения.
- Участие в Программах:
- Федеральная целевая программа «Экологическая безопасность России».
- Федеральная целевая программа «Перспективные информационные технологии».
- Программа глобального экологического фонда (ГЭФ) «Сохранение биоразнообразия».
- Региональные программы, заказчиками которых были: Правительство Ленинградской области; Министерство природных ресурсов Татарстана; Комитет природных ресурсов Башкортостан.

- Программах и проектах Государственного Комитета Рыболовства РФ по следующим подразделениям: СевЗапРыбВод; КамчатРыбВод; СахалинРыбВод; БалтРыбВод; КаспНИИРХ.
- Программа оценки воздействия на окружающую среду в Байкальском регионе и зоне БАМа.
- Программ Минэкономразвития по созданию свободной экономической зоны Кавминвод Ставропольский край и на Селигерской природной территории.
- Программа оптимизации региональной природопользовательской (охотохозяйственной, рыбохозяйственной) деятельности в Байкальском регионе.
- Проекты по обеспечению деятельности ЦНИЛ Главохота РФ.
- Проект Экономической комиссии ООН для Республики Беларусь (Создание геоинформационной аналитической системы для управления переводом ТЭК Республики на биотопливо).
- Разработка систем сертификации по ИСО 9001 и 14001 ОАО «Газпром», дочерними предприятиями «Газпрома»: Тюментрансгаз и Лентрансгаз.

1.5.2. Реализация принципов экономики знаний на основе БИИ

Эффективность современного бизнеса во многом определяется умелым использованием новейших информационных технологий и средств коммуникаций, сбора и обработки информации, принятия управленческих решений, политики продаж, проведения маркетинга и рекламы. При этом самой востребованной является информация, представленная в виде профилированных знаний экономического, экологического, социального характера, а также инвестиционной направленности. Достаточно понятно, что привлечение достоверных и актуальных знаний ускоряет процесс управления бизнесом, делает его эффективным и прибыльным, снижает потребность в физически производимом труде (а, значит и затраты на него), повышает конкурентоспособность, безопасность ведения и устойчивое развитие бизнеса.

К числу основных требований к информационному обеспечению бизнес-процессов относятся:

- получение решений системы в виде готовых рекомендаций, альтернатив, сценариев развития бизнес-процессов и бизнес-ситуаций;
- возможность работы с информацией, представленной в разнообразных числовых и лингвистических массивах данных и знаний;

- гибкость и динамичность информационных систем, под чем понимается способность прикладных программного обеспечения легко перестраиваться в условиях меняющихся условий бизнеса, требований и ситуаций;

- просчет рисков решений и управление ими;
- разработка инвестиционных предложений и проектов;
- быстрое решение производственных, маркетинговых и снабженческих, инвестиционных и кадровых задач;
- достоверная оценка влияющих факторов и ситуации в окружающей бизнес-среде;
- удобство общения с информационной системой, понятность выводов и рекомендаций системы;
- надежность, объективность, полнота бизнес-решений.

Все эти требования современного бизнеса определяют переход экономики постиндустриального общества к максимальному потреблению знаний, то есть формируют экономику, основанную на знаниях как производственном ресурсе, дополнительной мощной производительной силе, по определению [1] «экономику знаний». Концепция экономики знаний приведена в работах [1]. отечественных и зарубежных ученых.

Массовое востребование знаний создает условия для представления знаний в виде товарного продукта и для формирования рынка знаний. Все это определяет перспективу и необходимость создания и развития индустрии знаний, прежде всего профессиональных, что, в свою очередь, определяет необходимость разработки средств и технологий получения знаний, их преобразования, хранения, передачи и использования для решения различных прикладных задач. Как известно, такие технологии и средства носят название интеллектуальных [3–9]. На их основе строятся интеллектуальные системы, которые производят информационные объекты – «знания» в виде:

- решений, моделей, выводов, рекомендаций, аналитических оценок, интерпретаций ситуаций, оценок степени влияния, зависимости и проявления свойств объектов; динамики и тенденций факторов; прогнозов; критериев; утверждений; мнений;
- обоснований, когнитивных графических образов; и других видов знаний.

Как видно из этого неполного списка, все эти виды продуктов информационных систем активно востребованы и используются в совре-

менном бизнесе. И чем мощнее, распределеннее и многопрофильнее бизнес, тем значительнее доля этих продуктов присутствует в нем.

По данным аналитических обзоров [1] доля этой продукции может составлять до 90% стоимости основного товарного продукта бизнеса и в будущем будет повышаться.

Согласно экономической теории [1], оценка стоимости интеллектуальной составляющей в себестоимости продукции может быть определена по формуле:

$$S_{int} = S_{market} - S_{trade}, \quad (59)$$

где S_{int} – стоимость интеллектуальной составляющей,

S_{market} – рыночная стоимость основной товарной продукции,

S_{trade} – стоимость материальных, энергетических и трудовых ресурсов (физического труда), затраченных на производства данного вида продукции.

При конкретной продаже и установленных фактических затратах формула (59) может быть использована для определения фактической стоимости интеллектуальных ресурсов.

Так называемая «добавленная» стоимость может быть интерпретирована как знание или предвидение рыночной ситуации, при которой будут получена планируемая прибыль.

Таким образом при продаже товара на самом деле присутствует интеллектуальный ресурс, который либо в прошлом (в виде накопленного опыта или приобретенных знаний и деловой информации) был оплачен, либо в настоящем времени должен быть приобретен и оплачен. Однако, с другой стороны, не любая информация может считаться знаниями, востребованными и эффективными для бизнеса и соответственно иметь право на коммерциализацию.

Для производства прикладных знаний, реально используемых в бизнес-процессах и повышающих капитализацию бизнеса и его эффективность требуются надежные и актуальные знания. То есть к знаниям, используемым в бизнесе предъявляются жесткие требования к их достоверности, интерпретированности и скорости их получения. В реальных условиях бизнеса в проектных решениях и задачах, связанных с организацией бизнеса в будущих периодах, а также быстро меняющихся условиях существующего рынка возникают риски и не-

определенности в определении условий продаж, конкретных процессов, востребованности продукта и так далее, и с ними риски задач ошибочных оценок и решений. Поэтому для бизнес-задач в условиях рисков и неопределенностей информации (практически всех современных бизнес-задач) необходимы специальные методы и средства, которые удовлетворяют данным и выше перечисленным требованиям.

Имеющаяся во всех бизнес-задачах неопределенность информации о достоверности, полноте, четкости наборов данных прикладной задачи, надежности информационных источников как функциональных преобразователей данных, условиях и факторах окружающей среды и их влиянии на бизнес приводят к рискам оценки, необходимости многоальтернативных решений, недостоверности получаемых решений.

Поэтому для получения бизнес-решений прежде всего необходим математический аппарат, предназначенный для снижения неопределенности и контроля рисков получаемых решений. И только после его создания или выбора можно приступать к выбору или разработке информационных технологий, а затем и систем. Уместно отметить, что и в экономических научных трудах отмечается корреляция кризиса экономической теории именно с отсутствием адекватного современным экономическим задачам математического аппарата. Объективной причиной неэффективности эконометрических методов, основанных на классической математической статистике, является невыполнимость ее основных требований, постулатов и допущений. Так, например, невыполнимость требования соблюдения одинаковых условий эксперимента для получения репрезентативной выборки в виду меняющихся реальных условий бизнеса и вообще в целом природных и социальных процессов

В данной статье рассматриваются методологические аспекты и опыт использования регуляризирующего байесовского подхода (РБП) для различных экономико-социальных и природоохранных задач. Данный математический аппарат, созданный в начале 90-х годов прошлого века, стал методологической основой целого ряда отечественных и международных проектов производственного и территориального развития, приведенных в [2–17].

Регуляризирующие и интегрирующие свойства РБП, а также его метрологическая обоснованность являются принципиальной основой для разработки интеллектуальных технологий и систем для бизнес-задач. Многомерная свертка информации, ее преобразований на многомодельной основе на базе гибких формализованных моделей с динамическими ограничениями (МДО) и шкал с динамическими огра-

нениями (ШДО) позволяют получать метрологически обоснованные бизнес-решения в виде рекомендаций, аналитических оценок, выводов, тенденций, многоальтернативных сценариев, гибких стратегий управления, комплексной оценки ситуаций и программ развития бизнеса, производственных и территориальных комплексов.

Математический аппарат РБП стал основой байесовских интеллектуальных технологий и систем, которые реализованы в виде управляющих, мониторинговых и аудлирующих систем разного профиля. Для быстрой разработки систем такого типа создана программная среда «Инфоаналитик», позволяющая в самые сжатые сроки реализовать интеллектуальную систему для решения конкретной прикладной задачи.

Важным аспектом при практическом использовании данных систем является подготовка пользователя из бизнес-среды. Для эффективного обучения пользователя без специального прохождения обучения созданы ряд интеллектуальных рабочих мест, которые позволяют пользователю регулярно обновлять свои профессиональные знания, получать новую деловую информацию, включать полученные знания в бизнес-процесс. На ИРМ легко реализуются корпоративные сети и сети профессионального взаимодействия специалистов.

Поскольку, как было определено в [1], прикладные знания в современном бизнесе имеют сущность товарного продукта со стоимостью $\{h(S_{int})\}$, очевидно наличие бизнес-знаний и использование их может существенно повысить капитализацию компании. Для определения стоимости нематериальных активов целесообразно рассмотреть жизненный цикл информационного объекта и товарного продукта «знание» и детализировать составляющие показателя для правильной оценки затрат при реализации, модернизации, создании нового бизнес-процесса в инновационных и инвестиционных проекта. Для их практической реализуемости в реальных условиях необходимо учитывать составляющие затрат всех фаз жизненного цикла интеллектуальных ресурсов данного бизнес-процесса. Иначе проект может быть отклонен инвестирующей организацией как нереализуемый или будет финансово убыточным.

Следуя принципам реализации современного производства, управление производством знаний можно вести на основе международных производственных стандартов серий ISO 9000, ISO 14000, OHSAS 18000 и других.

Управление по стандарту ISO 9000 предполагает контроль качества производства и продукции знаний и может быть реализовано по классической схеме Деминга, предусматривающей этапы планирова-

ния, реализации, контроля и корректировки производственных мероприятий. Это, по сути, направлено на метрологию и повышение качества получаемых знаний. Введение таких схем контроля, с одной стороны, удорожает производство знаний, но с другой стороны повышает их ценность для бизнеса, а, следовательно, и их стоимость, что должно мотивировать их производителей.

Экологический стандарт ISO 14000 применительно к производству знаний определяет их полезность и, так сказать «экологичность» информационной среды, ее нравственность и гуманность.

Стандарт OHSAS 18000 направлен на подготовку персонала как активного и важнейшего фактора производства знаний.

Важным этапом для реализации производства знаний и управления им является создание модели процесса производства.

Любой бизнес-процесс может быть представлен как последовательность этапов маркетинга, планирования бизнеса, сбора ресурсов, производства, хранения, транспортировки, продаж продукции или ее утилизации.

В соответствии с этим процесс производства бизнес-знаний может быть представлен аналогично основному бизнес-процессу той же последовательностью этапов. На каждом из них создается определенная стоимость соответствующих этапу знаний.

Так на этапе маркетинга создается модель существующей и прогнозируемой рыночной ситуации на основе сбора данных о конъюнктуре, конкурентах, субститутах, потребителях, предпочтениях, насыщении, географии и других рыночных показателях. Применяя методологию байесовских интеллектуальных технологий (БИТ) в качестве знания этого этапа можно представить БИТ-модель маркетинга данного бизнес-процесса в виде многомерной байесовской свертки перечисленных выше показателей.

Для управления проектами при их реализации список должен быть дополнен временными показателями, отражающими бизнес-процесс. Для этого можно использовать методологию БИТ для управления проектами в условиях неопределенности и рисков.

Риски в бизнес-процессах – это неопределенность знаний. Используя соответствующие метрологически обоснованные знания, можно снять или снизить риски. Поэтому знания действительно истинный ресурс бизнеса, который может быть активизирован средствами интеллектуальных систем.

Знания всегда ограничены во времени и в пространстве, составляя в этих средах определенный информационный компакт. Но их можно непрерывно обновлять, дополнять и накапливать. Таким образом компакт знаний динамичен, что требует адекватного математического представления. Методология РБП основана на концепции динамического компакта пространства решений, что позволяет создавать модели объектов с динамическими ограничениями (МДО). Для бизнес-процессов это означает, что при получении новой информации модель бизнес-знаний и бизнес-процесса будет меняться, адаптируясь к новой ситуации самостоятельно. Это касается как сфер производства знаний, так и их использования.

Формализованная метрологическая основа БИТ-технологий позволяет определять степень практической приемлемости получаемых оценок, прогнозов, рекомендаций, обеспечивает возможность управления качеством решений на основе согласованного решения прямой и обратной задач метрологии БИИ и БИТ. Эти же свойства БИТ-технологий создают принципиальную основу и методическую базу для задач инвестирования и оценки риска новых проектов, реконструкции предприятий, расширения производств, строительства сооружений, комплексного и гармоничного развития территорий

1.5.3. Решение социально-экономических задач на основе байесовских интеллектуальных технологий

Современные социально-экономические процессы формируются в условиях уникальности ситуаций; активно влияющей, меняющейся окружающей среды; наличия региональной специфики; недостаточности, неточности, нечеткости информации, что обуславливает нестационарный характер этих процессов, а экономические модели основываются на постулатах о детерминированности факторов или о возможности получения экспериментальной информации о них в любом объеме. Ни того, ни другого на практике нет.

Несоответствие реальных процессов и их эконометрических моделей делает невозможным получение адекватных оценок и эффективных управленческих решений. Поэтому возникает сомнение в работоспособности, эффективности и пригодности эконометрических моделей для решения современных задач экономики. Однако, дело здесь не в самих методах, а в правомочности их использования в конкретных прикладных задачах.

Постановки современных социально-экономических задач (СЭЗ) обуславливают решение их в условиях значительной неопределенности информации. Действительно, решения, получаемые с использованием информационных технологий и систем, необходимы в сложных задачах, при недостаточности информации, когда человек не может самостоятельно найти их решения

Поэтому, очевидной является рекомендация использовать для решения таких задач соответствующие методы, ориентированные на применение в условиях неопределенности в широком смысле этого слова. Неопределенность может выражаться в недостаточной структурированности решаемых задач, так называемой их слабой обусловленности, неточности, неполноте, нечеткости информации, разнесенности информации во времени и в пространстве.

Одним из подходов, ориентированным на такие задачи, является байесовский регуляризирующий подход (РБП). Технологии на его основе (байесовские интеллектуальные технологии (БИТ)) нашли применение в решении широкого круга социально-экономических задач, таких как SWOT-аналитическая оценка состояния региональной и муниципальной экономики и социальной сферы, раз работка паспортов территорий, определение инвестиционных потенциалов территорий, кадастровое обоснование ресурсного потенциала, рациональное использование природных ресурсов и охрана окружающей среды, ландшафтное и территориальное планирование, организация туристско-рекреационных комплексов, разработка инвестиционных проектов и программ и других задачах. Опыт использования РБП показал эффективность его применения в задачах развития сельских территорий, что особенно важно в условиях активизации программ, реализации ряда правительственных распоряжений [1], направленных на устойчивое развитие сельского хозяйства.

Программный комплекс «Инфоаналитик» применен для решения основных социально-экономических задач, таких как задачи обеспечения устойчивого развития региональных и муниципальных территорий, задачи оценки ресурсного потенциала территорий, создание и введение в практику системы сбалансированных индексов и индикаторов показателей развития территорий (в целом и по отраслям хозяйствования); определение тенденций и динамики развития природных и социально-экономических процессов, нормирование показателей, индексов и индикаторов развития, аудит состояния территории по системе выбранных сбалансированных показателей и критериев развития, определение ин-

дексов и индикаторов, проведение аудита хозяйственной и управленческой деятельности на территории, получение совокупности оценок рисков и потенциалов территории, формирование инвестиционных планов и мероприятий, разработки конкретных инновационных и инвестиционных программ, создание центров по подготовке квалифицированных производственных и управленческих кадров и других производственных, управленческих, научных задачах и СЭЗ.

Опыт применения БИТ позволил разработать методическую основу решения этих задач, которая была реализована в виде пилотируемых и адаптируемых к прикладным задачам технологий. Основная концептуальная идея состоит в разработке на базе БИТ и платформ «Экоаналитик» и «Инфоаналитик» развивающихся технологий[5] для конкретных прикладных задач.

Концепция и подробное описание этого типа технологий приведено в [3].

1.5.4. Моделирование банковского сектора и показателей экономических систем ЕАЭС

Комплексная макро модель для экономических систем ЕАЭС включает в себя основные показатели экономических систем стран ЕАЭС. Она представлена на нижеследующих рисунках. В качестве нескольких примеров моделей банковского сектора ЕАЭС приведены модели в виде мягких функций авторегрессии для факторов «Международные резервы» и «Инфляция».

Метрология и оценка риска моделей:

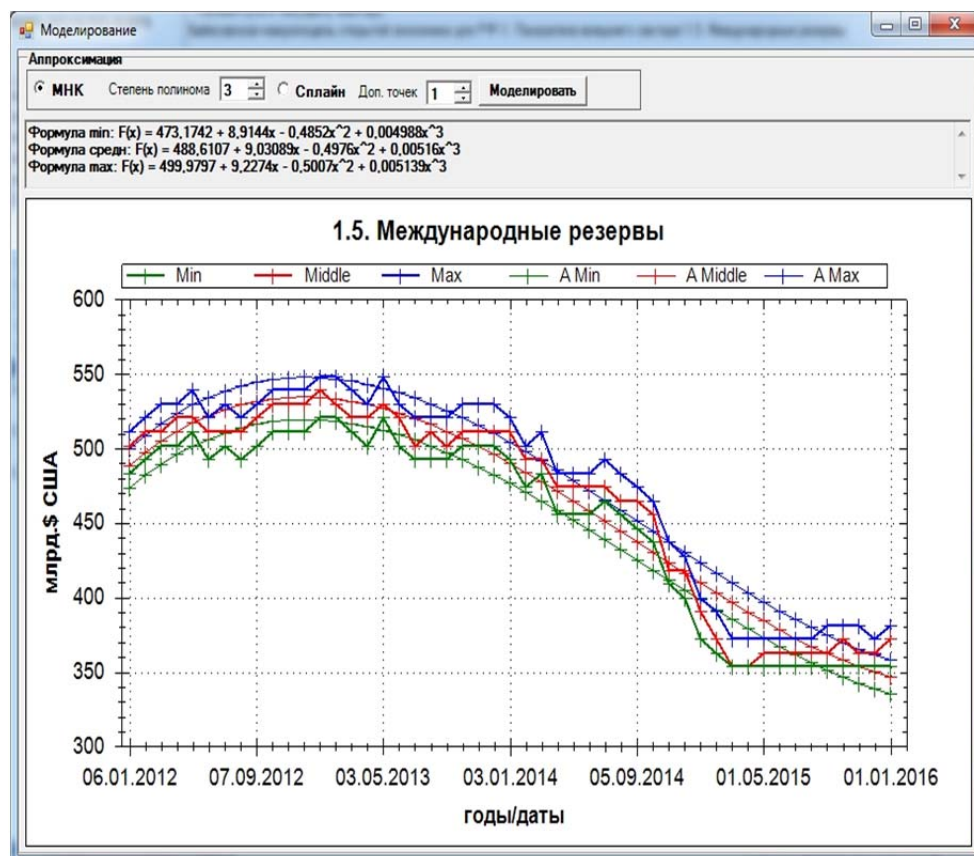
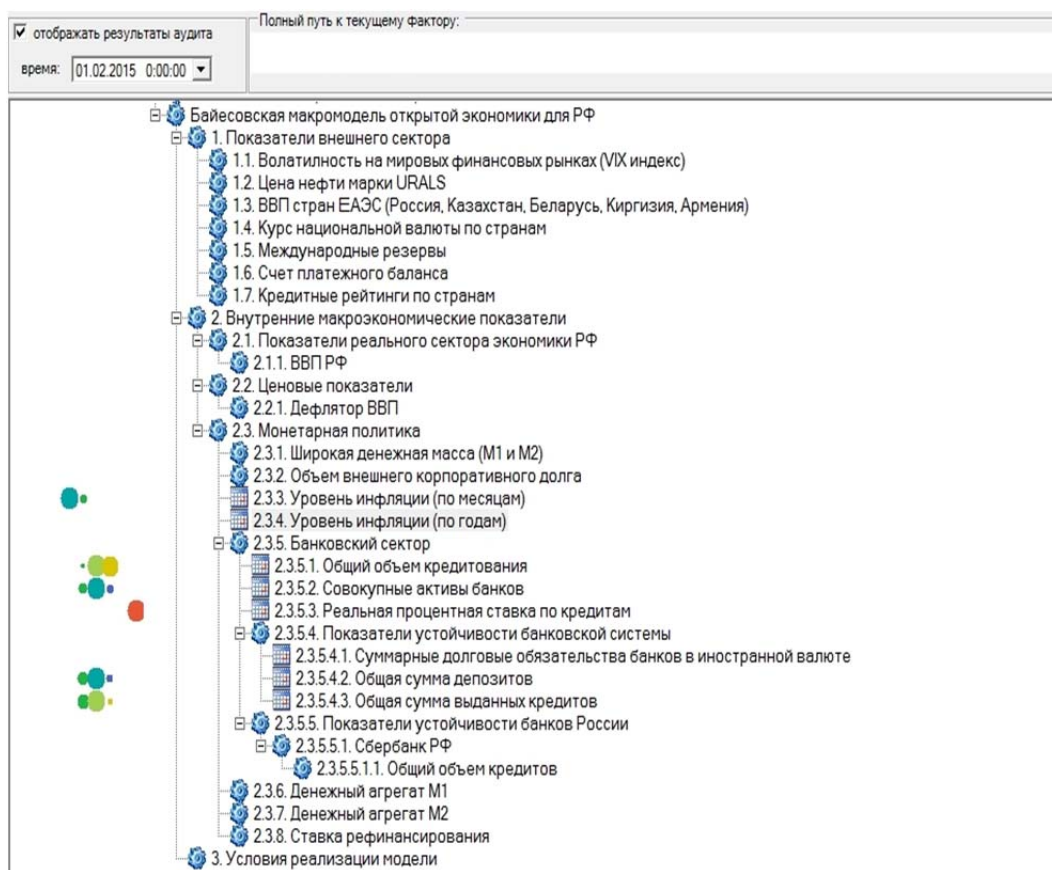
Метрологические характеристики и оценка качества модели верхнего уровня:

- Точность 2,9%.
- Достоверность 0,248.
- Надёжность 0,465.
- Риск применения модели 75,2%.

Метрологические характеристики и оценка качества наиболее вероятной модели:

- Точность 2,9%.
- Достоверность 0,528.
- Надёжность 0,465.
- Риск применения модели 47,2%.

Глава 1. Мягкие подходы к измерениям и управлению сложными системами. Регуляризирующий байесовский подход



Метрологические характеристики и оценка качества модели нижнего уровня:

- Точность 2,9%.
- Достоверность 0,223.
- Надёжность 0,465.
- Риск применения модели 77,7%.

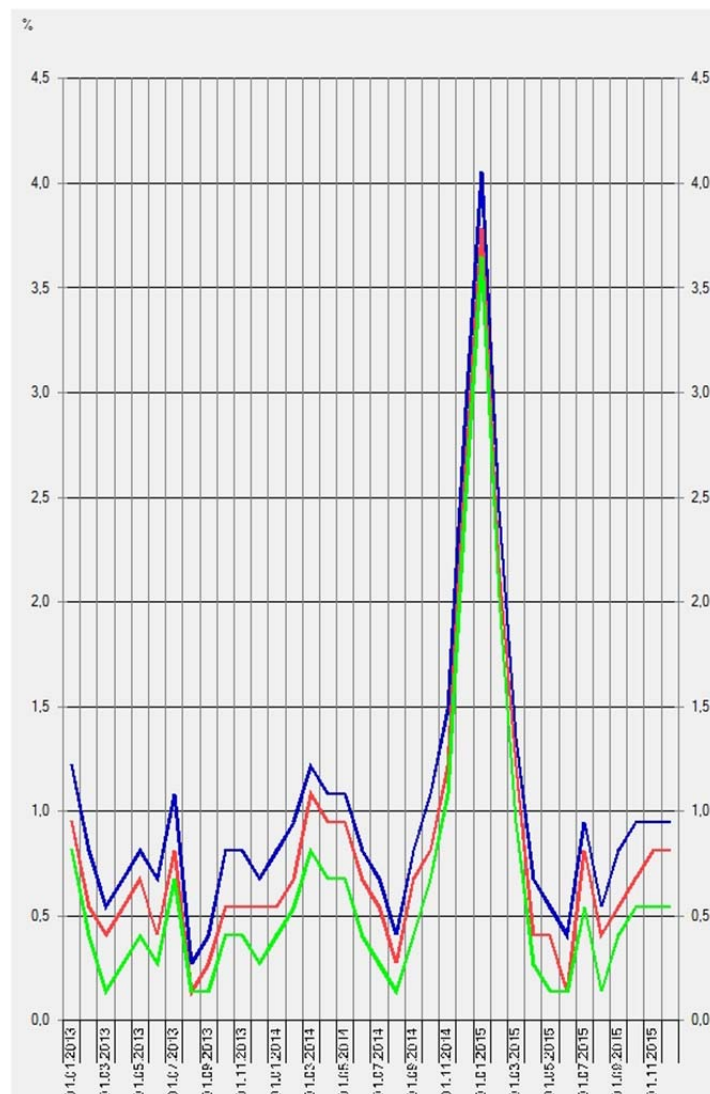
Мягкая модель в виде нечеткой функции авторегрессии для фактора «Международные резервы»:

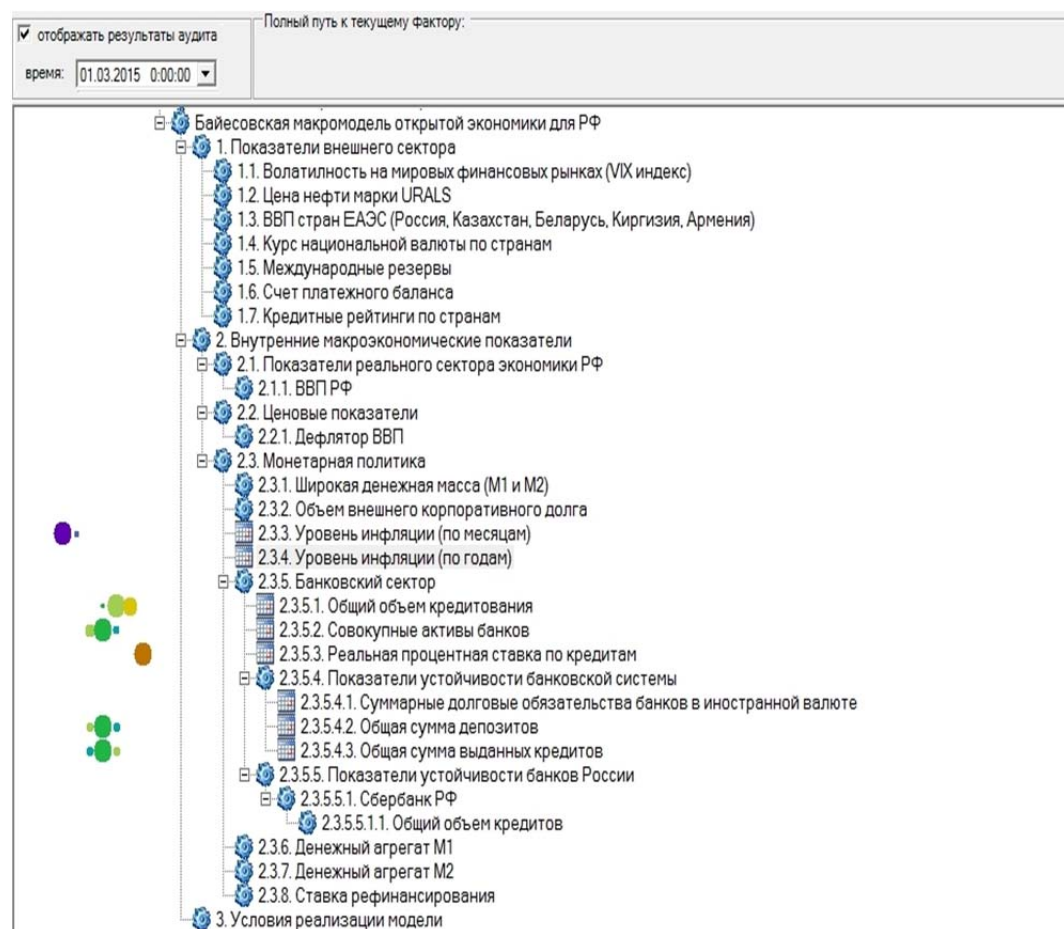
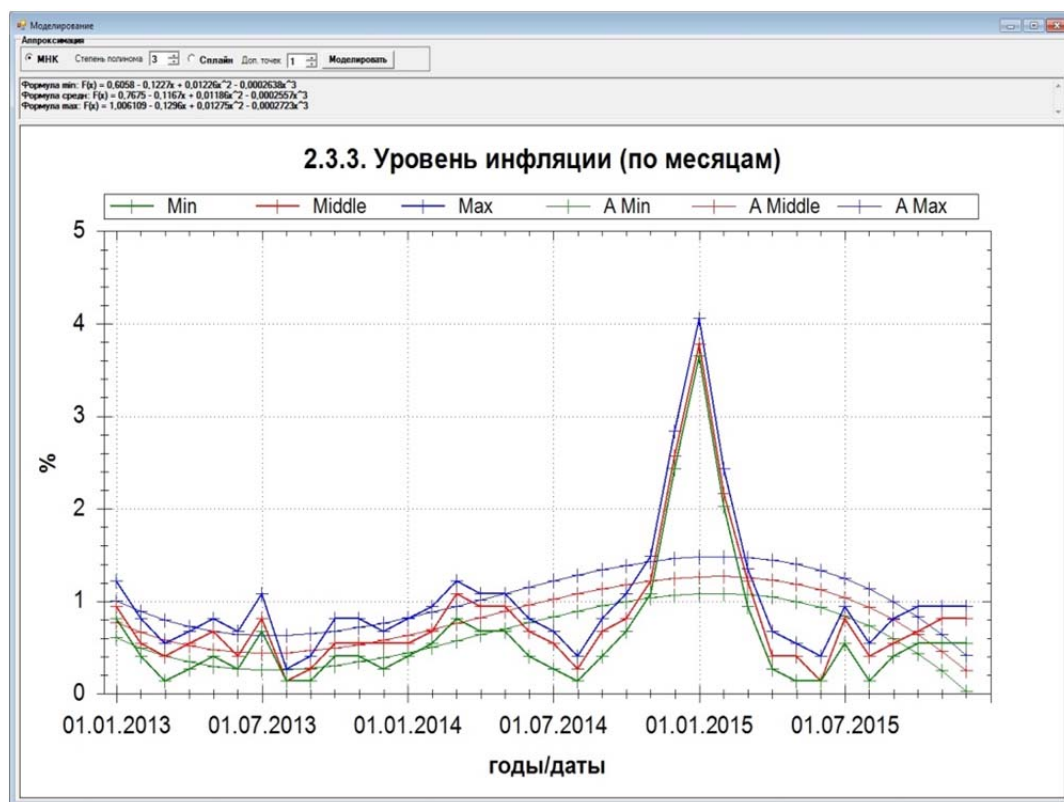
$$F(x) = 473,1742 + 8,9144x - 0,4852x^2 + 0,004988x^3$$

$$F(x) = 488,6107 + 9,03089x - 0,4976x^2 + 0,00516x^3$$

$$F(x) = 499,9797 + 9,2274x - 0,5007x^2 + 0,005139x^3$$

Моделирование показателя «Инфляция (по месяцам)»





Все решения получены в прикладных системах на основе платформы «Инфоаналитик» [12], представляющей собой программный комплекс для быстрой разработки приложения по оцениванию, измерению, математической обработке информации, генерации управленческих решений в условиях неопределенности.

1.5.5. Методологические аспекты создания интеллектуальных энергетических сетей (smart grid)

Мягкие технологии на основе РБП находят применение в задачах создания интеллектуальных энергетических сетей (smart grid).

Система управления технологическими активами, системы энергоаудита и управления энергосервисом, системы поддержки принятия решений для тарификации и разработки инвестиционных программ на основе программного продукта «Инфоаналитик» могут служить основой для разработки и внедрения систем энергоменеджмента (серия стандартов ИСО 50001) на основе повышения эффективности и качества управления энергосистемами.

Разработаны концепции создания интеллектуальных энергетических сетей (smart grid) и ее реализации для электроэнергетических сетей и сетей использования сжиженного природного газа.

Опыт создания этих систем мониторинга, контроля и управления на объектах энергетики и ЖКХ показывает, что они реально позволяют обеспечить:

- контроль экономичности работы оборудования с выводом информации о фактических расходах и затратах на энергоносители;
- снижение эксплуатационных затрат путем уменьшения общего времени на ремонт за счет минимизации времени поиска неисправности, увеличение межремонтного периода за счет своевременного проведения профилактических работ;
- контроль расходов энергоносителей на собственные нужды;
- выявление отклонений фактических размеров нагрузок от расчетных для учета в местных нормативах;
- формирование информационной базы для сравнительного анализа работы различных объектов и подключенных к ним систем;
- повышение безопасности работы оборудования.

Все это говорит о необходимости в целях повышения энергоэффективности и энергобезопасности в РФ, создания единой информационно-аналитической системы энергоэффективности и энергобезопасности.

Процессы нарастания дефицита топливно-энергетических ресурсов, безудержный рост цен, катастрофический износ основных фондов приводят к тому, что органы управления различных уровней во все большей степени нуждаются в полной и достоверной информации. При этом в едином информационном блоке должны быть представлены с одной стороны натуральные показатели энергобалансов различных уровней, а с другой их финансово-экономические показатели. Особенно важно включение в эту систему предприятий энергетики, которые в большей мере обеспечивают энергобезопасность на разных уровнях. Использование современных интеллектуальных информационных технологий значительно повышает эффективность оценки информации и качество принятия управленческих решений.

Создаваемая информационно-аналитическая система является распределенной и многоуровневой. Это федеральный, региональный, районный (городской) и уровень предприятий (организаций).

Структура ИАС должна обеспечивать:

- распределенный и параллельный принцип сбора, хранения и обработки информации;
- распределение функций системы по иерархическим уровням в соответствии с классами решаемых задач;
- единые методики и технические средства сбора, хранения, обработки и передачи информации;
- создание единой информационной и вычислительной среды для решения задач энерго- и ресурсосбережения, тарификации и лимитирования всех видов энергоресурсов;
- возможность оперативного и централизованного наблюдения за всеми потребителями энергоресурсов.

Общую архитектуру ИАС можно отнести к классу «клиент-сервер». Система федерального уровня предназначена для хранения интегрированной информации о статистических показателях потребления ТЭР в региональном и отраслевом разрезах. Система оснащается подсистемой поддержки принятия решений, которая разрабатывается на основе блока аналитической и математической обработки исходных данных. Пользователями системы являются руководители правительства и министерств, отвечающих за работу энергетики и жилищно-коммунальных комплексов. Информационная поддержка процесса принятия управленческих решений осуществляется на основе информации, поставляемой в систему с систем регионального уровня.

Системы регионального уровня предназначены для хранения интегрированной и детализированной информации о статистических показателях потребления энергоресурсов. Системы регионального уровня также оснащаются подсистемой поддержки принятия решений.

Системы уровня района (города) располагаются при администрациях муниципальных образований и обеспечивают сбор информации от предприятий и организаций и передачу ее в региональную систему.

Системы уровня предприятия (организации) располагаются на вычислительных средствах отдельных предприятий (организаций) и предназначены для хранения информации о статистических показателях потребления ТЭР, состоянии материальных активов и оборудовании. Пользователями системы уровня предприятия являются руководители и служащие энергетических служб предприятий.

Структура ИАС содержит:

- головной федеральный сервер, связанный информационными потоками с региональными системами;
- сервера региональных систем, предназначенные для сбора информации с систем муниципальных образований;
- сервера муниципальных систем, предназначенные для сбора первичной информации с предприятий и организаций.

В состав ИАС входят:

- распределенное информационное хранилище (совокупность баз данных);
- коммуникационная подсистема (совокупность коммуникационных средств);
- программные средства аналитической обработки данных и знаний.

Распределенное информационное хранилище на любом уровне иерархии системы включает сервер баз данных и собственно базы данных, построенные по реляционной архитектуре. Сервер баз данных обеспечивает: возможность многопользовательного доступа к информации, развитие функции поиска, сортировки и предоставления информации, создание витрин данных. Все информационно-поисковые, аналитические и служебные программные приложения (в том числе и работающие на удалении) единообразно взаимодействуют с сервером баз данных в качестве его клиентов. По характеру хранящейся информации банк данных ИАС образуетея двумя различными подсистемами:

- документно-ориентированной;
- фактографической (сигнальной).

Документно-ориентированная подсистема оперирует информацией нормативно-правового и нормативно-методического содержания; фактографическая – информацией, получаемой в результате сбора и последующей обработки первичных данных о статистических показателях потребления ТЭР. Фактографическая подсистема представляет многофакторную модель всей совокупности производителей и потребителей энергоресурсов. Документно-ориентированная подсистема предназначена для нормативно-методической поддержки различных информационно-аналитических функций ИАС нормативно-правового обоснования управленческих решений.

Коммуникационная система ИАС образуется: средствами Internet, обеспечивающими непосредственную работу клиентов с различными частями информационного хранилища, а также репликацию данных между Web-средствами E-mail, применением промежуточных носителей для переноса файлов – пополнений между физически удалёнными частями распределённого информационного хранилища.

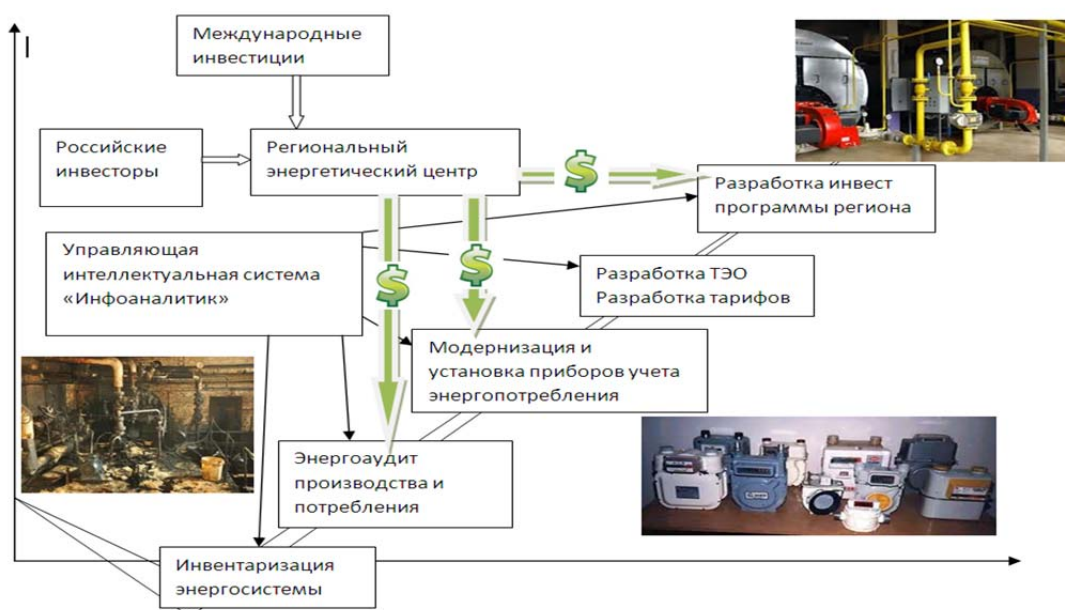
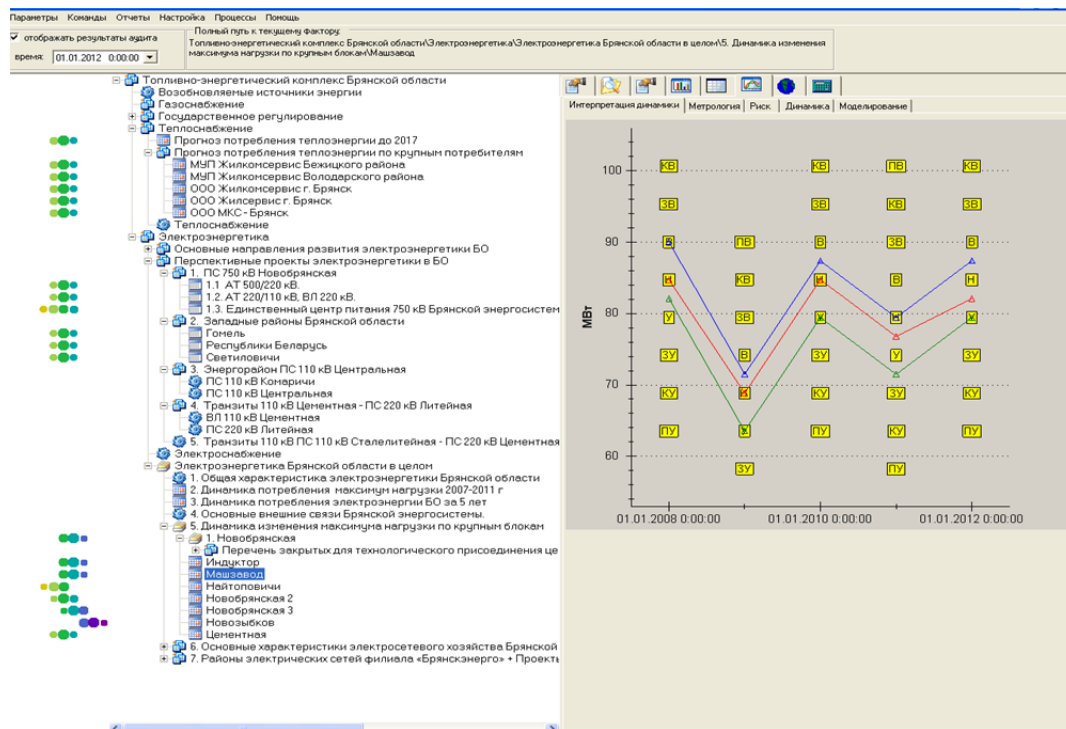


Рис. 26.

В качестве программного комплекса для многоуровневой федеральной информационно-аналитической системы может быть предложен программный комплекс «Инфоаналитик». Этот комплекс обеспечивает решение задач мониторинга, аудита, моделирования и управления сложными объектами в условиях неопределённости. В

качестве объектов в данном случае могут выступать энергетические и жилищно-коммунальные комплексы, предприятия по разработке и поставке местных видов топлива, экологические системы.

В качестве базовой платформы используется программная среда «Инфоаналитик». Она включает в себя систему создания моделей объектов анализа, подсистему сбора и передачи данных, систему интеллектуальной обработки и анализа данных, библиотеку байесовской математической статистики, базы моделей, систему визуального и когнитивного отображения введенной и обработанной информации, систему выгрузки данных и аналитической информации, подсистему работы с ГИС.



При работе с системой можно выделить следующие модули:

- Модуль «Работа с факторами»

Предназначен для построения моделей объектов и свойств, работы с классификационными шкалами и описаниями факторов – характеристик модели.

- Модуль «Характеристики фактора»

Используется для задания типа фактора, типа влияния фактора, единицы измерения, коэффициента масштабирования, корреляционного периода и регулируемости фактора.

- Модуль «Работа со шкалами»

Предназначен для создания шкал ШДО оценки свойств объектов по различным типам информационных Блок «Табличный ввод данных».

Используется для ввода данных в табличном виде. Данные представлять собой числовое значение, экспертную оценку, шкальное данное или текстовую информацию.

- Модуль «Динамическая шкала»

Служит для отображения динамики и тенденций развития фактора.

- Модуль «Метрология»

Данный модуль обеспечивает оценку характеристик метрологических комплексов (точности, надежности, достоверности, риска, энтропии получаемых решений).

- Модуль «Рекомендации»

Предназначен для анализа фактора – интерпретации ситуации и выдачи рекомендаций по управлению фактором и влияющими факторами.

- Модуль «Моделирование»

Если характеристики модели связаны между собой функциональными (аналитическими) зависимостями, модуль определяет аналитический вид зависимости в виде кусочно-линейной модели, полиномиальной или иной модели.

Список литературы

1. *Прокопчина С.В.* Системы мониторинга и поддержки принятия решений в природоохроне и природопользовании. СПб., 1998. 11 с.
2. *Недосекин Д.Д., Прокопчина С.В., Чернявский Е.А.* Информационные технологии интеллектуализации измерительных процессах. СПб.: Энергоатомиздат, 1995. 187 с.
3. *Прокопчина С.В.* Организация измерительных процессов в условиях неопределенности. Регуляризирующий байесовский подход. СПб. Сборник докладов Международной конференции по мягким вычислениям и измерениям SCM-98, 22–26 июня 1998. Т. 1. С. 30–44.
4. *Прокопчина С.В.* Концепция байесовской интеллектуализации измерений в задачах мониторинга сложных объектов // Новости искусственного интеллекта.
5. *Прокопчина С.В.* Экспертная система АССИСТЕНТ для интеллектуальных байесовских измерений вероятностных характеристик // Измерительная техника. 1992. № 8. С. 11.
6. *Прокопчина С.В.* Байесовские интеллектуальные технологии для аудита и управления сложными объектами в условиях значитель-

ной неопределенности / Сб. докладов IX Международной конференции по мягким вычислениям и измерениям SCM-2005.

7. *Прокопчина С.В.* Синтез развивающихся информационных технологий управления на транспорте // Научно-технический журнал «Технико-технологические проблемы сервиса». № 2. 2007. С. 119–130.

8. *Прокопчина С.В., Федичкин А.И.* Применение байесовских интеллектуальных технологий для оценки и прогнозирования состояний сложных объектов // Научно-технический журнал «Технико-технологические проблемы сервиса». № 2. 2007. С. 108–118.

9. *Прокопчина С.В.* Байесовские интеллектуальные технологии и системы для комплексного управления территориями. Сборник докладов первого Международного инвестиционного форума «Инновации и инвестиции в интеллектуальные технологии управления предприятиями и территориальным развитием». Санкт-Петербург, 25–27 июня 2007. С. 117–137.

10. *Прокопчина С.В.* Решение социально-экономических задач на основе байесовских интеллектуальных технологий. СПб. Сборник докладов Международной конференции по мягким вычислениям и измерениям SCM-2011, 23–25 июня. Т. 1.

11. Интеллектуальное Рабочее Место менеджера предприятия (свидетельство Федеральной Службы по интеллектуальной собственности, патентам и торговым знакам об официальной регистрации программы для ЭВМ № 2006610206 от 21.03.2006).

12. Интеллектуальное Рабочее Место специалиста по качеству (свидетельство Роспатента об официальной регистрации программы для ЭВМ № 2004610246 от 19.02.2004).

13. Интеллектуальное Рабочее Место финансового директора (свидетельство Федеральной Службы по интеллектуальной собственности, патентам и торговым знакам об официальной регистрации программы для ЭВМ № 2005611423 от 08.08.2005).

14. Инфоаналитик (свидетельство Федеральной Службы по интеллектуальной собственности, патентам и торговым знакам об официальной регистрации программы для ЭВМ № 2004611741 от 12.08.2004).

15. Интеллектуальный Тренинговый Комплекс для специалистов предприятия (свидетельство Роспатента об официальной регистрации программы для ЭВМ № 2006610210 от 21.03.2006).

16. *Прокопчина С.В.* Методологические аспекты теории мягких измерений Сб. докл. Междунар. конф. «SCM-2009». СПб. С. 49–63.

17. Управление в условиях неопределенности. Под ред. С.В. Прокопчиной, М.Ю. Шестопалова. СПб.: Изд. ЛЭТИ, 2014. 375 с.

18. Опыт решения социально-экономических задач на основе БИТ / под ред. Прокопчиной С.В. М., 2014. 447 с.

19. *Прокопчина С.В.* Принципы мягкого аудита и мягкого управления сложными системами. М., 2010.

20. Интеллектуальное Рабочее Место специалиста по управлению территорией (свидетельство Федеральной Службы по интеллектуальной собственности, патентам и торговым знакам об официальной регистрации программы для ЭВМ № 2006610207 от 21.03.2006.

21. *Прокопчина С.В.* Синтез развивающихся информационных технологий управления на транспорте // Научно-технический журнал «Технико-технологические проблемы сервиса». № 2. 2007. С. 119–130.

22. *Прокопчина С.В., Федичкин А.И.* Применение байесовских интеллектуальных технологий для оценки и прогнозирования состояний сложных объектов // Научно-технический журнал «Технико-технологические проблемы сервиса». № 2. 2007. С. 108–115.

23. *Прокопчина С.В.* Байесовские интеллектуальные технологии и системы для комплексного управления территориями. Сборник докладов первого Международного инвестиционного форума «Инновации и инвестиции в интеллектуальные технологии управления предприятиями и территориальным развитием». Санкт-Петербург, 25–27 июня 2007. С. 117–137.

24. *Prokopchina S.V., Averkin A.N.* The short essay of the soft measurement concept. SPb.: Hidrometeoizdat, 1997. 38 p.

25. *Prokopchina S.V.* Project Management Under Conditions of a Priority Uncertainty on the base of Regularizing Bayesian Approach Proceedings of I international Conference on Soft computing Technologies in Economy. Pp. 165–171.

26. *Kupriyanov M.S., Prokopchina S.V.* Creation of algorithmic models to control self-regularizing redundant systems. ICAFS-2014, Proceedings of 11-th International Conference on Application of Fuzzy Systems and Soft Computing, Paris, 2–3, September, 2014.

27. *Prokopchina S.V.* The Principles of the creation of the developing information technologies in the basis of Regularizing Bayesian Approach. ICAFS-2014, Proceedings of 11-th International Conference on Application of Fuzzy Systems and Soft Computing, Paris, 2–3, September, 2014.

ГЛАВА 2

ОТ НЕЧЕТКИХ МНОЖЕСТВ К МЯГКИМ ОЦЕНКАМ И СИНЕРГЕТИЧЕСКОМУ ИСКУССТВЕННОМУ ИНТЕЛЛЕКТУ

В.Б. Тарасов

Введение

В 1965 г. в международном журнале «Information and Control» была опубликована новаторская работа знаменитого американского ученого Лотфи Заде «Нечеткие множества» (Fuzzy Sets). Суть оригинальных воззрений Л. Заде состояла в том, что реальные человеческие рассуждения, опирающиеся на естественный язык, не могут быть описаны в рамках традиционных математических формализмов. Введение *нечетких множеств* – классов с неточно определенными границами, описываемых функциями принадлежности (обобщающими характеристические функции обычных множеств) – обеспечило основу для развития более гибкого подхода к анализу рассуждений и моделированию сложных, трудноформализуемых, гуманистических систем. В исходной статье [1] Л. Заде дал определение нечеткого множества A как функции $\mu_A: U \rightarrow [0,1]$, изложил различные варианты задания бинарных операций пересечения и объединения для нечетких множеств, ввел унарные операции дополнения, концентрирования, растяжения, отношение включения для нечетких множеств, предложил понятия нечетких отношений и их проекций, рассмотрел четкие множества уровня нечеткого множества. Им были также предложены принцип обобщения и нечеткие множества типа 2 (с нечеткими функциями принадлежности), введены понятия нечетких ограничений и правил трансляции для нечетких предложений, приведены примеры, иллюстрирующие предложенный подход.

Эта работа всколыхнула научный мир и вызвала к жизни целый ряд прорывных направлений в области философской логики и прикладной математики, системного анализа и теории принятия решений,

информатики и искусственного интеллекта, многие из которых были открыты в последующих статьях самого Л. Заде [2–8]. В частности, сюда относятся нечеткая логика [2] и приближенные рассуждения [3], теория лингвистических переменных [4] и вычисления со словами [5], теория возможности [6] и теория гранулирования информации [7], мягкие вычисления [8] и вычислительный интеллект.

Особо хотелось бы отметить, что Л. Заде был сформулирован общий принцип *несовместимости* (по мере роста сложности исследуемой системы наша способность точно описывать ее поведение уменьшается вплоть до порога, за которым точность и релевантность становятся практически взаимоисключающими характеристиками) и предложен единый подход к моделированию *неточности* границ и *неопределенности* на основе нечетких множеств (функций распределения возможности) [1, 4, 6]. Взаимосвязи между различными видами неопределенности и нечеткости исследовались многими авторами. Например, в [9] А.Н. Борисовым и др. была дана классификация видов неопределенности в задачах управления. В ней различаются *неизвестность*, *неполнота* и *недостоверность* информации, причем выделяются ситуации физической (объективной, стохастической) и лингвистической (субъективной, нестохастической) недостоверности.

Схожая классификация для задач проектирования была предложена В.Б. Тарасовым (см. [10, 11]). В этой работе были выделены ключевые свойства инженерных знаний на начальных этапах проектирования и разработаны варианты их формализации. В частности, отмечены: а) *неполнота*, фрагментарность исходных данных и суждений, обусловленные недостаточным, неполным знанием характеристик новых материалов, особенностей технологических процессов, условий работы новых конструкций; б) *нечеткость* представлений о взаимосвязях между проектными параметрами, связанная с отсутствием четких аналитических зависимостей между ними или вычислительной сложностью; в) принципиальные *ограничения по точности* задания как количественных параметров, так и, в особенности, качественных факторов решения (для таких факторов, как качество, простота, ремонтпригодность инженерных конструкций естественные физические шкалы вообще отсутствуют; их учет всегда зависит от опыта и интуиции специалиста).

В свою очередь, Д. Дюбуа и А. Прад [12] также проводили главное разграничение между факторами *неточностью* и *неопределенностью*. Неточность относится к содержанию имеющейся информации, а неопределенность – к ее истинности, понимаемой либо в смысле со-

ответствия реальной действительности, либо как правдивости эксперта. При этом ими был разработан единый подход к моделированию этих двух видов факторов на основе мер (квазимер) возможности и необходимости, а также интерпретации функции распределения возможности как нечеткого множества.

Пожалуй, наиболее основательный содержательный анализ различных свойств человеческих знаний, тесно связанных с особенностями естественного языка, провел А.С. Нариньяни. В 1982-м году им был введен в научный обиход термин «НЕ-факторы знаний» для обозначения комплекса свойств (*недоопределенность, неточность, неполнота, некорректность* и др.), характерных для реальных знаний, но плохо представленных в формальных системах [13, 14]. Подход А.С. Нариньяни опирается на тщательный анализ функционирования НЕ-факторов в языке; в его трактовке, инженерия знаний – это переход от доформального исследования объекта моделирования к его формальной модели. При этом наиболее исследованным НЕ-фактором у него остается *недоопределенность*, который формально описывается с помощью недоопределенных множеств [13, 14].

Примерно в то же время: Ж. Дельгранде и Дж. Милопулос [15], исходя из анализа различных случаев противоречий, неполноты, незамкнутости в базах знаний, предложили свой вариант концепции НЕ-факторов знаний, используя для их обозначения составную аббревиатуру Im-In-Un's. Обзор результатов логического подхода к анализу НЕ-факторов можно найти в [16], а первые попытки развития гибридного лингвологического подхода к моделированию НЕ-факторов изложены в [17, 18].

2.1. НЕ-факторы, мягкие оценки и инженерия мнений

В работе [8] Л. Заде описал концепцию *мягких вычислений*, объединяющих подходы нечеткой логики, нейроинтеллекта и эволюционного моделирования. При этом обычно нечеткие продукционные модели реализуются в обучаемых нейронных сетях, оптимизация которых происходит с помощью генетических алгоритмов. Затем в [19] А.Н. Аверкин и С.В. Прокопчина ввели понятие *мягких измерений* (в этом контексте ими рассматривались, в частности, пути объединения вероятностных и нечетких измерений). В настоящей работе обсуждаются варианты интерпретации и формального представления *мягких оценок*, предложенных в [20].

Предварительно укажем на некоторые содержательные и формальные различия между измерением и оцениванием. По своей сути измерения носят чисто структурный, синтаксический характер, тогда как оценки всегда имеют семантико-прагматическое наполнение.

Как известно, классическое понятие шкалы в теории измерений определяется как гомоморфизм (т.е. однозначное отображение, сохраняющее порядок) из эмпирической системы с отношениями в числовую систему с отношениями. Таким образом, каждому отношению в исходной, плохо структурированной системе соответствует его аналог в другой, хорошо структурированной системе. Поэтому первая основная проблема теории измерений заключается в установлении формальных свойств эмпирических соотношений и операций и доказательстве их изоморфизма соответствующим образом выбранным отношениям и операциям над числами. Вторая основная проблема связана с определением типа шкалы, с помощью которой проводятся измерения.

Кроме того, по традиции, идущей еще от Р. Льюса и Дж. Галантера, классическая теория измерений развивалась на стыке с теорией вероятностей и математической статистикой. Поэтому измерение предполагало определение классического пространства с вероятностной (аддитивной) мерой. Однако, во многих реальных приложениях требование аддитивности не выполняется. Поэтому более адекватным аппаратом здесь видится одна из теорий расширений вероятностных мер, например, теория Демпстера-Шейфера, теория псевдомер на основе треугольных конорм и т.п.

В противовес гомоморфизмам теории измерений, теория оценивания обычно имеет дело с отношениями полиморфизма (т.е. с одно-многозначными и много-многозначными отображениями). В частности, речь идет о нечетких отображениях, когда некоторому объекту $u \in U$ ставится в соответствие (с разными степенями истинности) нечеткое множество объектов $B \subset Z$, описываемое функцией принадлежности μ_B . Примером подобных неклассических определений шкалы как отношения полиморфизма служит нечеткое соответствие между лингвистическими значениями и их числовыми оценками в случае лингвистической переменной. Поэтому важная специфическая особенность теории оценивания заключается в формировании так называемых обобщенных шкал [21].

В свое время наличие пар антонимов (высокий – низкий, острый – тупой, добрый – злой и т.п.) в естественных языках всех типов навело

Ч. Осгуда на мысль о том, что они являются образующими модели мира у человека. Так появилось представление о многомерном психосемантическом пространстве, в которое проецируются знания. Согласно идее семантического дифференциала (различителя), понятия в этом пространстве расположены не хаотично, а формируют кластеры по принципу семантической близости. Экспериментальные данные показывают, что кластеры образуют не разбиение психосемантического пространства, а его многослойное покрытие.

Это наблюдение привело Д.А. Поспелова к убеждению о необходимости создания математической теории градуального «сходства – различия», т.е. теории толерантности, опирающейся на нечеткие отношения с переменным уровнем рефлексивности и симметричности. В основе этой теории лежит отказ от классических принципов жесткой различимости и принадлежности. Как следствие отказа от принципа различимости, было сформулировано представление об обобщенной шкале. В отличие от обычных шкал, где каждой точке соотносится один-единственный объект, на обобщенных шкалах любой точке может с разными степенями соответствовать множество объектов.

В работе [21] Д.А. Поспеловым среди всех оппозиционных (полярных) шкал были выделены «серые» и «черно-белые» шкалы», а также исследованы их основные характеристики. При оценивании на «серых» шкалах переход от свойства A^+ (например, «большой») некоторого объекта (к примеру, «дом») к противоположному свойству A^- («малый») происходит плавно, постепенно. Подобные шкалы удовлетворяют условиям: а) взаимной компенсации между свойствами A^+ и A^- (чем в большей степени проявляется A^+ , тем в меньшей степени проявляется A^- , и наоборот); б) наличия нейтральной точки A^0 , интерпретируемой как точка наибольшего противоречия, в которой оба свойства присутствуют в равной степени $A^+ = A^-$ (т.е. в данном примере имеется в виду дом, при оценке размеров которого в равной степени можно считать его как большим, так и малым). Обычно для «серых» шкал полагается, что свойства A^+ и A^- связаны между собой логическим отрицанием, т.е. $A^- = \neg A^+$.

Для метрических полярных шкал, когда можно определить расстояния между любыми двумя точками шкалы, первое условие выражает идею дополненности (сумма расстояний от любой точки шкалы до ее концов равна общей длине шкалы). Второе условие свидетельствует о наличии срединной точки, где расстояния до левого и правого концов шкалы равны между собой.

Напротив, в случае «черно-белых» шкал компенсации противоположных свойств не происходит: при уменьшении степени проявления A^+ степень проявления A^- не возрастает. Здесь принцип дополненности не работает: по мере удаления от краев шкалы оценки обоих свойств как A^+ , так и A^- становятся все более неопределенными. Более того, срединная точка подобной шкалы есть точка разрыва, в которой нет ни того, ни другого свойства $\neg (A^+ \vee A^-)$, т.е. наблюдается максимальная неопределенность.

В результате, было обосновано утверждение, что шкалы – образующие модели мира подчас не являются классическими оппозиционными шкалами. Соответственно, в статье [21] были введены круговые (или кольцевые) шкалы и предложены двухосновные оценки объектов на таких шкалах, которые выражаются парами вида $\{(x, y) \mid x \in X, y \in Y\}$, где X и Y – упорядоченные множества.

Жесткими оценками называются такие оценки, которые принимают значения на булевых шкалах. Тогда под *мягкой оценкой* в широком смысле понимается любая многозначная, нечеткая или лингвистическая оценка, пронизанная НЕ-факторами. Мягкая оценка в узком смысле содержит НЕ-факторы нескольких (по крайней мере, двух) типов [20]. Тогда мягкая оценка на полярной шкале может быть представлена в виде [17] $\mu = (\mu_{A^+}, \mu_{A^-})$, где μ_{A^+} и μ_{A^-} – нечеткие множества, характеризующие лингвистические оценки одновременного проявления положительного свойства A^+ и отрицательного свойства A^- .

В общем случае, мягкие лингвистические оценки на шкалах описываются в виде $SE=M$ (C есть tA), где t – модификатор, характеризующий *неточную* лингвистическую оценку по критерию C (например, модификатор «очень» в оценке величины «очень большая»), а M – модальность, выражающая показатель *неопределенности* этой оценки (например, «возможно», «вероятно», «желательно» и пр.). В случае полярных шкал имеем противоречивые мягкие оценки вида $SE^* = M$ (C есть t_1A^+ И t_2A^-), где A^+ и A^- – положительная и отрицательная оценки, а t_1 и t_2 – соответствующие модификаторы.

Подобные мягкие оценки можно положить в основу инженерии индивидуальных и коллективных мнений, которые занимают промежуточное положение между знанием и незнанием и не могут быть общезначимыми. В отличие от инженерии знаний в инженерии мнений центральное место занимают процедуры описания НЕ-факторов и их взаимопревращений.

Ниже изложим варианты формализации ряда типичных НЕ-факторов в русле предлагаемого здесь лингвологического подхода.

2.2. Лингвологический подход к формализации НЕ-факторов: принципы, этапы, модели

Лингвологический подход к формализации НЕ-факторов опирается на принципы: а) логической локальности [2], плюрализма и релятивизма; б) тесной взаимосвязи между логикой и онтологией; в) относительности, частичности семантической категории «истинности» (с чем связана необходимость учета ее градаций, многозначности и разрывов, а также дополнения истинности различными прагматическими категориями, такими как «полезность», «релевантность», «устойчивость» и пр.); г) множественности адекватных и взаимопереходящих логических (лингвологических) операций; д) единого алгебраического представления логических структур с помощью расширенных логических и логико-семиотических матриц [18]. Он включает пять основных этапов: 1) анализ сходства и различия между различными НЕ-факторами в естественном языке; спецификация выбранного НЕ-фактора (группы НЕ-факторов); 2) выбор базовой модели описания НЕ-фактора (например, интервала, вероятности или нечеткого множества); 3) подбор соответствующей логики; 4) идентификация классов операций логики (в первую очередь, классов отрицаний), зависящих от онтологии; 5) построение локальной модели ситуации с использованием логико-семиотической матрицы.

Ряд аспектов организации и проведения лингвистических исследований на первом этапе описан в [14]. Остановимся подробнее на втором и третьем этапах.

Фундаментальными законами классической логики являются закон полноты и закон непротиворечивости (в логической семантике им соответствуют закон бивалентности и закон однозначности). Отбрасывая последовательно эти законы, получаем следующие два класса многозначных логик: 1) *параполные* (или *частичные*) логики, для которых не выполняется аксиома полноты (закон исключенного третьего); 2) *пара-непротиворечивые* логики – не выполняется закон непротиворечия. Естественно положить результаты исследований данных многозначных логик в основу моделирования таких НЕ-факторов как *неполнота* и

противоречивость. В свою очередь, рассмотрение таких частичных логик, для которых не выполняется закон *отрицания отрицания*, означает переход к интуиционистским логикам, а одновременное нарушение свойств полноты и непротиворечивости характерно для *паранормальных* логик, примерами которых служат *нечеткие* логики.

В случае многозначных и нечетких логик информацию о классах и ролях НЕ-факторов несут промежуточные логические значения, лежащие между истиной и ложью, а также соответствующие этим логикам алгебры (например, алгебра Клини, алгебра Гейтинга, алгебра Чэна и др.).

Предложение 1. Минимальным классом многозначных логик для учета и моделирования НЕ-факторов является класс трехзначных логик.

Примеры. В трехзначной логике Я. Лукасевича третье, промежуточное значение 0,5 между истиной 1 и ложью 0 интерпретируется либо как *безразличие*, либо как *случайность*. В трехзначной частичной логике С. Клини величина 0,5 означает *неизвестность* (*неполноту информации*), а в трехзначной логике Д.А. Бочвара значение 0,5 понимается как бессмыслица (абсурд). Наконец, в трехзначной паранепротиворечивой логике Асеньо значение 0,5 отождествляется с *противоречием*.

Предложение 2. Минимальным классом многозначных логик для учета и моделирования взаимодействия НЕ-факторов является класс четырехзначных логик.

Предложение 2*. Класс четырехзначных логик есть минимальный класс *гибридных логик*, т.е. логик, обеспечивающих совместный учет НЕ-факторов различных типов.

Примеры. Двумя хорошо известными четырехзначными логиками, позволяющими работать с НЕ-факторами, являются логика Н. Белнапа (где помимо классических значений истинности Т и F введены значения В и N – противоречие и полная неопределенность соответственно) и логика аргументации В.К. Финна (когда ситуация противоречия в споре, связанная с (почти) одинаковым числом аргументов «за» и «против», вырождается в ситуацию неопределенности при отсутствии аргументов как «за», так и «против»).

2.2.1. Моделирование нечеткости

Основные пути моделирования *нечеткости* (этап 2 лингвологического подхода) можно систематизировать с помощью следующих критериев [10]:

а) общий подход к представлению нечеткости (нечеткие множества, описываемые функциями принадлежности; семейства обычных множеств; вероятностные множества; двухосновные нечеткие множества и т.п.);

б) область определения функции принадлежности (стандартные нечеткие множества, нечеткие множества уровня, нечеткие числа, и пр.);

в) область значений функции принадлежности ($[0,1]$ -нечеткие множества, $[-1,+1]$ -нечеткие множества, L -нечеткие множества, интервальнозначные нечеткие множества, нечеткие множества типа 2 и пр.);

г) тип соответствия между областью определения и областью значений функции принадлежности (взаимно однозначное, один-ко-многим, и пр.).

Стандартная нечеткая логика Л. Заде представляет собой прямое обобщение классической двузначной логики. Большинство ее известных расширений, например, логика, строящаяся на основе предложенных в 1967 г. Дж. Гогеном L -нечетких множеств [22], которые принимают значения в некоторой решетке, опираются на все те же принципы бивалентности и однозначности классической логической семантики.

Нетрудно понять, что для многих реальных ситуаций, характеризующихся неопределенностью или противоречивостью, необходимо выбрать другие, неклассические базовые семантики. Например, если требуется расширить частичную или паранепротиворечивую логику, то за основу надо взять трехзначную семантику, а при фаззификации минимальных гибридных логик (учитывающих НЕ-факторы по крайней мере двух типов) следует опираться на исходную четырехзначную семантику Данна-Белнапа.

Для расширения областей значений функций принадлежности могут использоваться произведения решеток. Так еще в 1974 г. А. де Люка и С. Термини [23] ввели $[0,1]^m$ -нечеткие множества с целью описания профилей объектов (или понятий), когда семейство нечетких множеств $\{A\}_m$ соответствует m свойствам объекта, а каждый элемент $x \in X$ характеризуется вектором значений принадлежности $(\mu_1(x), \dots, (\mu_m(x)))$. Затем Л. Кочи ввел понятие векторнозначного нечеткого множества как отображения $\mu: X \rightarrow F_1 \times F_m$, где F_i , $i = 1, \dots, m$ – ограниченные линейно упорядоченные множества. Очень интересным обобщением L -нечеткого множества является предложенное А. Кофманом гетерогенное нечеткое множество (см. [10]), где различные элементы множества X принимают свои значения из различных решеток.

Среди альтернативных (гибридных) подходов следует упомянуть ряд моделей двухосновных нечетких множеств. Так в середине 80-х годов К. Атанасов ввел так называемые *интуиционистские нечеткие множества* [24], описываемые парой функций принадлежности μ и непринадлежности ν соответственно, $A = \{(x, \mu_A(x), \nu_A(x))\}$. Свой отказ от принципа однозначности он обосновал, исходя из важной идеи интуиционизма об относительной независимости утверждения и отрицания. Аналогичным образом, П. Рен [25] предложил обобщенные нечеткие множества для работы с неполной информацией. В свою очередь, Д. Дюбуа и А. Прад [26] также ввели двухосновные нечеткие множества, где элементы пары функций принадлежности выражали соответственно степени возможности и необходимости принадлежности элемента плохо определенному четкому множеству. Еще одна разновидность двухосновных нечетких множеств получается в результате обобщения процедуры экспертного оценивания на полярных шкалах [17].

Тем не менее, общая теоретическая модель для формализации мягких оценок на шкалах и построения гибридных нечетких логик пока еще отсутствует. На наш взгляд, перспективы ее построения тесно связаны с аппаратом *произведений решеток* и, в частности, *бирешеток* (bilattices) [27].

2.2.2. Моделирование противоречивости

Пусть обобщенная лингвологическая оценка LLE на полярной шкале есть пара {утверждение, отрицание}, т.е. $LLE = (m(p), n(p))$, где m – операция утверждения, а n – операция отрицания (заданные функционально-аксиоматически [10, 17]). Далее для упрощения возьмем в качестве типичного оператора утверждения тождественное преобразование $m(x) = x$ и будем рассматривать пару $(x, n(x))$.

Тогда класс *показателей противоречивости* лингвологических оценок задается с помощью выражения $xTi(x)$, где T – произвольная треугольная норма (см., например, [10]), соответствующая лингвистической связке «И», а n – унарный оператор отрицания.

Введем отношение квазипорядка, индуцируемое оператором отрицания n (т.е. рефлексивное и транзитивное отношение) $\leq_n$ на L_n : $x \leq_n y \Leftrightarrow x T n(x) \leq y T n(y)$ и будем рассматривать квазиупорядоченное множество L_n .

Возьмем для определенности в качестве T наибольшую треугольную норму: $T = \wedge = \min$ и будем описывать противоречивые оценки

на полярных шкалах с помощью следующего отношения квазипорядка: $x \leq_n y \Leftrightarrow x \wedge n(x) \leq y \wedge n(y)$. В частности, оценка x на полярной шкале является более противоречивой, чем оценка y , если x и $n(x)$ лежат между y и $n(y)$. Очевидно, что в случае $L = [-1, +1]$ максимально противоречивой оценке отвечает условие $x=n(x)=0$, а совершенно непротиворечивой оценкой будет оценка, соответствующая краям шкалы: $x = +1 \Rightarrow n(x) = -1$ и наоборот $x = -1 \Rightarrow n(x) = +1$.

2.2.3. Моделирование неточности

Пусть неточные оценки на шкалах представляют собой подмножества A решетки L , т.е. $A \subseteq L$. Тогда класс мер неточности оценок на шкалах можно определить в виде функционалов $I: 2^L \rightarrow \mathbb{R}^+$, таких, что: (i) $I(A) = 0$ тогда и только тогда, когда подмножество A является одноточечным, т.е. $A = \{x\}$; (ii) $\forall A, B \in 2^L, A \subseteq B \Rightarrow I(A) \leq I(B)$.

2.3. От семейств НЕ-факторов к гибридным моделям и синергетическому ИИ

В этом разделе обсуждаются взаимосвязи между моделированием НЕ-факторов, мягкими вычислениями и созданием интеллектуальных систем новых поколений, воплощающих идеи и принципы синергетики в русле искусственного интеллекта.

Термин «синергетика» происходит от слова «синергия», означающего *совместное действие, сотрудничество*. По мнению «отца синергетики» Г. Хакена, такое название хорошо подходит для современной теории сложных самоорганизующихся систем по двум причинам: а) исследуются совместные действия многих элементов развивающейся системы; б) для отыскания общих принципов самоорганизации требуется объединение усилий представителей различных дисциплин. В последние годы синергетика нередко понимается как наука о взаимодействиях [28], на основе которых раскрываются общие принципы самоорганизации, эволюции и кооперации сложных систем.

Впервые концепция *синергетического искусственного интеллекта* была представлена в [29]. В качестве объекта синергетического ИИ выступают сложные самоорганизующиеся интеллектуальные системы. Подобные системы обладают такими характеристиками как: а) наличие множества неоднородных компонентов n ($n \geq 2$); б) активность (целена-

правленность), автономность, кооперативное поведение компонентов; в) наличие множества различных взаимосвязей между компонентами, носящих семиотический характер; г) открытость, распределенность, сетевая организация; д) неустойчивость, адаптивность; е) высокий эволюционный потенциал, обучаемость. К числу важнейших принципов синергетического ИИ относятся: принцип инженерии взаимодействий; принцип дополнительности, предполагающий сочетание, гибридизацию различных моделей и методов описания в интересах взаимной компенсации их недостатков и усиления достоинств; принцип учета взаимодействия множества НЕ-факторов как ключевого механизма самоорганизации и развития (так из синергетики известно, что *неустойчивость* есть причина самоорганизации, а *неравновесность* влечет за собой процессы упорядочения и реструктуризации. *Нелинейность* порождает новые стратегии интеграции, а *необратимость* определяет когерентность, взаимодействие компонентов в масштабе всей системы).

Поскольку в русле синергетического ИИ особую значимость приобретает изучение сочетаний НЕ-факторов различных классов, возникает проблема развития соответствующего математического инструментария. По нашему мнению, в его основе будут лежать диалоговые логики [30] и формальный аппарат бирешеток [27]. Понятие бирешетки, предложенное М. Гинсбергом [27], задается четверкой $BL = (L_1 \times L_2, \leq_1, \leq_2, f)$, где $(L_1, \leq_1)$ и $(L_2, \leq_2)$ – полные решетки, а f – унарная операция на X , такая что: 1) если $x \leq_1 y$, то $f(x) \geq_1 f(y)$; 2) если $x \leq_2 y$, то $f(x) \leq_2 f(y)$; 3) $f(f(x)) = x$. Часто предполагается, что $L_1 = L_2 = L$, т.е. вместо $L_1 \times L_2$ в определении бирешетки берется L^2 .

Синергетический потенциал теории бирешеток обусловлен тем, что ее аппарат базируется на установлении взаимосвязей между двумя различными отношениями порядка. Так у М. Гинсберга это осуществляется с помощью специальной унарной операции, называемой им отрицанием (хотя в строгом формальном смысле она есть гибридный отрицания и утверждения), а М. Фиттингом определяется специальный подкласс сплетенных бирешеток (interlaced bilattices) [31].

В заключение, следует отметить, что аппарат бирешеток (их свойства и операции) позволяет исследовать с единых позиций взаимосвязи как между различными видами нестандартных нечетких множеств (например, между интервальнозначными и интуиционистскими нечеткими множествами), так и между различными когнитивными и коммуникативными задачами, предполагающими использо-

вание мягких оценок (оценивание на полярных шкалах, аргументация, взаимодействие на основе протокола диалоговых игр и пр.).

Список литературы

1. *Zadeh L.A.* Fuzzy Sets // *Information and Control*. 1965. Vol. 8. Pp. 338–353.
2. *Bellman R., Zadeh L.A.* Local and Fuzzy Logics // *Modern Uses of Multiple-Valued Logics* / Ed. by J.M. Dunn and G. Epstein. – Dordrecht: D. Reidel, 1977. Pp. 105–165.
3. *Zadeh L.A.* A Theory of Approximate Reasoning // *Machine Intelligence* / Ed. by J. Hayes, D. Michie and L.I. Mikulich. – New York: Halstead Press, 1979. Pp. 149–194.
4. *Zadeh L.A.* The Concept of a Linguistic Variable and its Application to Approximate Reasoning. Parts 1 and 2 // *Information Sciences*. 1975. Vol. 8. Pp. 199–249, 301–357.
5. *Zadeh L.A.* Fuzzy Logic = Computing With Words // *IEEE Transactions on Fuzzy Systems*. 1996. Vol. 4. Pp. 103–111.
6. *Zadeh L.A.* Fuzzy Sets as a Basis for a Theory of Possibility // *Fuzzy Sets and Systems*. 1978. Vol. 1. Pp. 3–28.
7. *Zadeh L.A.* Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic // *Fuzzy Sets and Systems*. 1997. Vol. 90. Pp. 111–127.
8. *Zadeh L.A.* Fuzzy Logic, Neural Network and Soft Computing // *Communications of the ACM*. 1994. Vol. 37. № 3. Pp. 77–84.
9. Модели принятия решений на основе лингвистической переменной / А.Н. Борисов, А.В. Алексеев, О.А. Крумберг и др. Рига: Зинатне, 1982.
10. Нечеткие множества в моделях управления и искусственного интеллекта / А.Н. Аверкин, И.З. Батыршин, А.Ф. Блишун, В.Б. Силов, В.Б. Тарасов. М.: Наука, 1986.
11. *Тарасов В.Б., Чернышев А.П.* Оптимизация деятельности проектировщика на ранних стадиях разработки летательных аппаратов // *Труды ВНИИТЭ. Серия «Эргономика»*. Вып. 36. Проблемы эргономического обеспечения проектирования технических средств. М.: ВНИИТЭ, 1988. С. 111–120.
12. *Дюбуа Д., Прад А.* Теория возможностей. Приложения к представлению знаний в информатике: Пер. с франц. М.: Радио и связь, 1990.

13. *Нариньяни А.С.* Недоопределенные модели и операции с недоопределенными значениями. Препринт ВЦ СО АН СССР. № 400. Новосибирск: ВЦ СО АН СССР, 1982.

14. *Нариньяни А.С.* НЕ-факторы: неточность и недоопределенность – различие и взаимосвязь // Известия РАН: Теория и системы управления. 2000. № 5. С. 44–56.

15. *Delgrande J.P., Mylopoulos J.* Knowledge Representation: Features of Knowledge // Fundamentals of Artificial Intelligence / Ed. by W. Bibel and Ph. Jorrand. Berlin: Springer Verlag, 1986. Pp. 3–36.

16. *Вагин В.Н.* Знание в интеллектуальных системах // Новости искусственного интеллекта. 2002. № 6. С. 8–18.

17. *Тарасов В.Б.* Анализ и моделирование НЕ-факторов на полярных шкалах // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник трудов Международного научно-практического семинара (Коломна, 17–18 мая 2001 г.). М.: Наука. Физматлит, 2001. С. 65–71.

18. *Тарасов В.Б.* От алгебраической модели Тарского к логико-семиотической матрице // Труды 9-й национальной конференции по искусственному интеллекту КИИ-2004 (Тверь, 28 сентября – 2 октября 2004 г.). Т. 1. М.: Физматлит, 2004. С. 311–322.

19. *Аверкин А.Н., Прокопчина С.В.* Мягкие вычисления и измерения // Интеллектуальные системы (МГУ). 1997. Т. 2. Вып. 1–4. С. 93–114.

20. *Tarassov V.B.* General Approaches to the Modelling of Soft Estimates and Beliefs in Strategic Decision Engineering // Proceedings of 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002, Divnomorskoe, Russia, 5–10 September). Los Alamitos CA: IEEE, 2002. Pp. 45–49.

21. *Поспелов Д.А.* «Серые» и/или «черно-белые» // Прикладная эргономика. Специальный выпуск «Рефлексивные процессы». 1994. № 1. С. 29–33.

22. *Goguen J.* L-Fuzzy Sets // Journal of Mathematical Analysis and Applications. 1967. Vol. 18. Pp. 145–174.

23. *De Luca A., Termini S.* Entropy of L-Fuzzy Sets // Information and Control. 1974. Vol. 24. Pp. 55–73.

24. *Atanassov K.T.* Intuitionistic Fuzzy Sets // Fuzzy Sets and Systems. 1986. Vol. 20. P. 87–96.

25. *Ren P.* Generalized Fuzzy Sets and Representation of Incomplete Knowledge // Fuzzy Sets and Systems. 1990. Vol. 36. Pp. 91–96.

26. *Dubois D., Prade H.* Twofold Fuzzy Sets and Rough Sets // *Fuzzy Sets and Systems*. 1987. Vol. 23. Pp. 3–18.
27. *Ginsberg M.* Multi-Valued Logics: a Uniform Approach to Reasoning in AI // *Computer Intelligence*. 1988. Vol. 4. Pp. 256–316.
28. *Хакен Г.* Тайны природы. Синергетика: учение о взаимодействии. Москва-Ижевск: ИКИ, 2003.
29. *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М.: Эдиториал УРСС, 2002.
30. *Felscher W.* Dialogues as a Foundation for Intuitionistic Logic // *Handbook of Philosophical Logic*. Vol. III / Ed.by D. Gabbay and F. Guenther. Dordrecht: D. Reidel, 1986. Pp. 341–372.
31. *Fitting M.* Bilattices and the Theory of Truth // *Journal of Philosophical Logic*. 1989. Vol. 18. Pp. 225–256.

ГЛАВА 3

О МЕТОДАХ ПОСТРОЕНИЯ ГРАНУЛЯРНЫХ ЛОГИЧЕСКИХ ЗНАЧЕНИЙ И СТРУКТУР

В.Б. Тарасов

Введение

Построение гранулярных логических структур предполагает обращение к современным основаниям логики, в частности, к идеям и подходам *универсальной логики* [27, 31, 32]. Универсальная логика (УЛ) – это не новая логика, а скорее попытка построить общую теорию логик, рассматриваемых как математические (в частности, алгебраические, геометрические, топологические) структуры. Главный инициатор развития УЛЖ.-И. Безье полагает, что она призвана играть роль, аналогичную роли универсальной алгебры при изучении различных алгебраических структур. Среди родоначальников УЛ следует указать таких выдающихся ученых как А. Тарский, А. Линденбаум, С. Яськовский, Р. Вуйцицкий, А.И. Мальцев и др.

Причина возникновения УЛ – реакция на логический плюрализм, появление в последнее время сотен новых логик, что влечет за собой потребность в их систематизации и упорядочения. В этом плане выделяются наиболее общие категории и единый инструментарий логики: логическое следование и оператор присоединения следствий, логическая система, логическое (истинностное) значение, логическая матрица, логическая семантика, логическое пространство, аналитические таблицы и пр.

Создание гранулярных логик означает как развитие теории грануляции [23], так и обобщение вышеуказанных базовых логических категорий. Главными объектами теории грануляции являются гранулы, уровни, гранулярные структуры, например, разбиения, иерархии, и соотношения между ними. Основными проблемами теории грануляции в логике являются: 1) определение общих принципов и подходов логиче-

ской грануляции; 2) создание, интерпретация и представление логических гранул; 3) построение моделей логических гранулярных структур; 4) анализ взаимосвязей между логическими гранулами; 5) анализ количественных характеристик гранул и процесса грануляции.

В настоящей работе основное внимание уделяется описанию и анализу ряда ключевых методов и инструментов логической грануляции: обобщенных значений истинности и нестандартных логических семантик, логических решеток, произведений логик и бирешеток, обобщенных матриц и логических пространств.

3.1. Гранулярные значения истинности и нестандартные семантики

3.1.1. Классические интерпретации значений истинности

Идея истинности высказывания известна в логике с незапамятных времен. В частности, ее разрабатывали Сократ, Аристотель, Парменид, позднее Фома Аквинский, Г. Лейбниц, И. Кант, Г.В. Гегель, а также основатель баденской школы неокантианцев В. Виндельбанд в русле базисной триады ценностей «Истина – Добро – Красота».

В явном виде понятие истинностного значения ввел выдающийся логик и философ конца XIX-начала XX веков Г. Фреге в своих статьях «Функция и понятие» и «О смысле и значении» (см. [12]). В философии языка Фреге высказывания трактуются как специальный вид имен, значениями которых являются два истинностных значения: истина или ложь. При этом истинность и ложность не являются имманентными свойствами высказываний, а представляют два первичных, наиболее фундаментальных объекта – «истину» (*das Wahre*) и «ложь» (*das Falsche*). Затем, «с легкой руки» А. Черча, значения истинности стали называть «абстрактными объектами».

В результате, логика получает онтологическое обоснование и характеризуется у Фреге как *«наука о наиболее общих законах бытия истины»* [12]. Близкие воззрения на предмет логики высказывал также Я. Лукасевич, который определял ее как *науку об истинностных (логических) значениях*: «Все истинные высказывания обозначают один и тот же объект, а именно истину, и все ложные высказывания обозначают один и тот же объект, а именно ложь. Я рассматриваю истину и ложь как единичные (*singular*) объекты... Онтологически аналогом истины

является бытие, а лжи – небытие. Объекты, обозначаемые высказываниями, называются логическими значениями... Логика есть наука об объектах особого рода, а именно, *наука о логических значениях*» [29].

В свою очередь, эпистемологический подход связан с интерпретацией истинности как некоторой оценки. При этом истина теряет свой абсолютный онтологический характер и обретает связь с абстрактными эпистемическими ситуациями (в смысле Дж. Данна) [18]. В этом русле истинность понимается как свойство, которое сохраняется при переходе от посылок к заключению в корректном рассуждении.

Таким образом, фундаментальные логические законы непосредственно связаны с онтологическими условиями, накладываемыми на данную логику. Например, «закон исключенного третьего» выражает постулат Фреге о существовании в точности двух истинностных значений – «истина» и «ложь». Двумя главными принципами семантики Фреге являются принцип бивалентности и принцип однозначности.

Пусть имеется множество высказываний P . В соответствии с принципом бивалентности, всякое высказывание $p \in P$ принимает одно из двух истинностных значений: $T(p) \vee F(p)$, а согласно принципу однозначности любое из этих двух значений есть одноточечное множество (синглетон) $T(p) = \{T\}$, $F(p) = \{F\}$ (это оговорено в вышеприведенном определении истины и лжи как сингулярных объектов по Я. Лукасевичу [29]).

Отказ от принципа бивалентности привел к появлению трехзначных, четырехзначных и других многозначных логик, развитию паранеполных логик с частичными значениями истинностями.

В то время как принцип бивалентности многократно подвергался критике, начиная с работ самого Аристотеля (в связи с опровержением принципа логического фатализма), Л. Брауэра, Я. Лукасевича [29], Н.А. Васильева [1] и др., принцип однозначности оспаривался куда реже. Среди логических результатов, связанных с отказом от принципа однозначности, следует указать, в первую очередь, воображаемую логику Н.А. Васильева [1], дискурсивную логику С. Яськовского, паранепротиворечивую логику Да Косты, логику парадоксов Г. Приста.

Классические значения истинности образуют двухэлементную булеву алгебру с операциями пересечения $\wedge$ и объединения $\vee$. Эта булева алгебра есть решетка с частичным порядком истинности $\leq: F \leq T = F \wedge T = F$. Взаимосвязь значений истинности с высказываниями обеспечивает функция истинностной оценки, т.е. функция из множества элементарных высказываний S_b $V_2 = \{T, F\}$:

$$v^2: S \rightarrow \{T, F\} \quad (1)$$

Здесь для любых p, q из S имеем $v^2(\sim p) = -v^2(p)$, $v^2(p \wedge q) = v^2(p) \wedge v^2(q)$, $v^2(p \vee q) = v^2(p) \vee v^2(q)$, $p \models q \Leftrightarrow \forall v^2, v^2(p) \leq v^2(q)$.

Обосновывая принятие истинностных значений в качестве экстенционалов предложений, Р. Карнап проводит строгую аналогию между объемами предикатов и истинностными значениями предложений [6]. При этом эквивалентность двух высказываний p и q означает совпадение их истинностных значений.

Представление значений истинности как абстрактных объектов отнюдь не является единственным. В последние три-четыре десятилетия появилось новые интерпретации истинностных значений, в частности:

- абстрактные сущности, рассматриваемые как классы эквивалентности высказываний;

- объекты, выражающие неопределенность или размытость понятий;
- значения, характеризующие степень истинности суждений;
- значения, передающие информацию о некотором предложении.

Для иллюстрации эквивалентностной трактовки истинностных значений определим для $V_2 = \{T, F\}$ m -кратное прямое произведение $V_2^m = V_2 \times V_2 \times \dots \times V_2$. Поскольку V_2 есть двухэлементное множество, то число элементов V_2^m равно 2^m . Любая Т-Ф-последовательность представляет собой мультимножество классических значений истинности, где функция кратности характеризует число вхождений Т и F. Таким образом, имеется интерпретация любых истинностных значений (натуральных чисел или дробей) в терминах ТкF как кратности Т или F, деленной на длину всей Т-Ф-последовательности. Например, значение истинности 0 понимается как «однородная F-последовательность», в которой все вхождения есть F, а значение $1/3$ может интерпретироваться последовательностью {Т, F, F}. Такая интерпретация, где в качестве истинностных значений выступают подмножества множества Т-Ф-последовательностей, была предложена А.С. Карпенко (см. [7]) и названа *фактор-семантикой*.

Остальные трактовки истинностных значений естественным образом предполагают привлечение концепции грануляции информации Л. Заде [34] в область логических семантик, градуирование значений истинности и построение гранулярных и/или градуированных семантик [8–10].

3.1.2. Нестандартные логические семантики и обобщенные значения истинности

В последние годы постоянно растет число нетрадиционных логических семантик. Примерами нестандартных семантик служат: семантика Васильева; семантика Данна-Белнапа и тесно связанные с ней бирешеточные семантики; многомерные (векторные) семантики, в частности, семантики возможности и необходимости; нечеткие семантики; диалоговые и игровые семантики.

Логические модели грануляции опираются на многозначные и паранепротиворечивые логики. Одним из родоначальников этих логик был русский ученый Н.А. Васильев [1], который построил двухуровневую логическую структуру (эмпирическая логика и металогика), а также ввел многомерные (воображаемые) логики [1], (иллюстрацией трехмерных логических высказываний является треугольник Васильева – см. рис. 2).

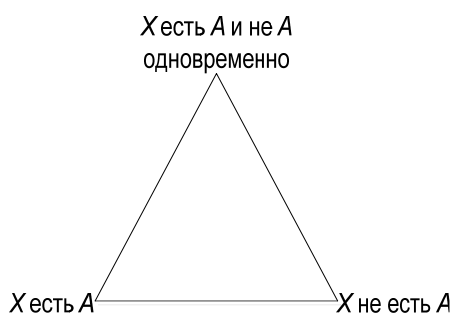


Рис. 1. Треугольник Васильева

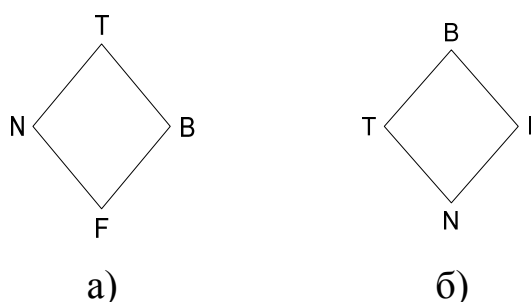


Рис. 2. Диаграммы Хассе для решеток Белнапа (а) и Скотта (б)

Одна из вершин этого треугольника, соответствующая составному противоречивому высказыванию « X есть A и не есть A одновременно», является примером логической гранулы. В то же время двухуровневая логическая структура может пониматься как ранний вариант логической грануляции, перехода от мелкозернистой к крупнозернистой семантике.

Основоположником неклассических логических семантик является Дж. Данн [18,19], который еще в середине 1960-х годов предложил новую стратегию построения логических семантик, в которой допускаются ситуации, когда высказывания являются одновременно и истинными, и ложными (отказ от принципа однозначности), или не являются ни истинными, ни ложными (отказ от принципа бивалентности). Первый из

таких нестандартных случаев называют «пресыщенной оценкой» (Glut), а второй – «истиннозначным провалом» (Gap) (см., например, [13]). Здесь истинность понимается как функция множества, т.е. ее значениями могут быть не только элементы множества V , но и любые его подмножества, включая пустое множество $\emptyset$, т.е. осуществляется переход от V к 2^V . Эти подмножества называются обобщенными истинностными значениями и вполне соответствуют трактовке истинности по Фреге-Черчу как обобщенного объекта. В простейшем случае для $|I|=2$ имеем $2^V = \{T, B, N, F\}$, где $B = \{T, F\}$, $N = \emptyset$ (логическая семантика Белнапа).

При построении своей четырехзначной логики Белнап рассмотрел два отношения частичного порядка: порядок истинности $\leq_v$ и порядок включения на множестве $\{\emptyset, \{T\}, \{F\}, \{T, F\}\}$. В первом случае имеем $F < N < T$ и $F < B < T$, а значения B и N , находящиеся между F и T , являются несравнимыми. Во втором случае получаем $\emptyset \subset \{T\} \subset \{T, F\}$ и $\emptyset \subset \{F\} \subset \{T, F\}$, т.е. T и F являются несравнимыми. Первый порядок задает полную решетку значений истинности, которая названа Белнапом логической решеткой $L4$, а второму порядку соответствует решетка Скотта $A4$ (которую в данном контексте следует информационной решеткой). Операция отрицания в логике Белнапа является составной, неоднородной и консервативной. Последнее означает, что оно сохраняет свойство классического отрицания $\neg T = F$ и $\neg F = T$ по порядку $\leq_v$, а неоднородность связана с рассмотрением отрицания по информационному порядку $\leq_i$ как тождественного преобразования, где значения B и N являются неподвижными точками. По сути, отрицание Белнапа есть обобщение отрицания Клини, а четырехзначная логика Белнапа является расширением трехзначной логики Клини.

Естественными расширениями этого подхода является использование нечетких подмножеств из $[0,1]^V$ (семантика Заде) [33] или их обобщений из L^V , где L – решетка (семантика Гогена) [22], а также описание значений истинности как термов лингвистической переменной ИСТИННОСТЬ [5]. В последнем случае устанавливается полиморфное соответствие между лингвистическими метками, образующих терм-множество понятия «Истинность» («истинно», «скорее истинно, чем ложно», «ложно», «абсолютно ложно» и пр.) и числовыми значениями истинности из интервала $[0,1]$, т.е. речь идет об отношениях «один-ко-многим».

Будем теперь применять в качестве общего способа построения логических гранул произведения логик, рассматриваемые как произведения решеток (по Г. Биркгофу). Следуя Н. Решеру [30], определим

модализированные значения истинности на основе $\{0, 1\} \times \{0, 1\}$. Пусть имеется пара алетических модальностей $\{L, Q\}$, где L – «необходимая», а Q – «случайная» (другой вариант – пара модальностей норм $\{P, Z\}$, P – «разрешенная», Z – «запрещенная») и пара классических значений истинности $\{T, F\}$. Тогда получаем обобщенные модализированные значения истинности (рис. 3).

Довольно часто встречаются ситуации, когда либо оценки истинности носят двумерный характер (например, теоретическая и эмпирическая истинность), либо рассматриваются комплексные логические значения с двумя порядками (истинность, информация) или (истинность, неточность). Тогда приходим к непосредственному расширению семантики Белнапа на основе аппарата бирешеток [20,21].

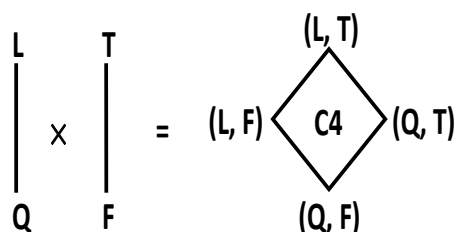


Рис. 3. Семантические гранулы как модализированные значения истинности

Прямое отношение к построению многомерных гранул имеют также концепция возможных миров [26], игровые [25] и диалоговые семантики [28, 11]. Так в [28] П. Лоренцен рассмотрел процесс определения истинности предложения как игру между двумя агентами (ролями) – пропонентом и оппонентом. Установлению истинности соответствует проигрыш в этой игре оппонента, когда он не может совершить никакого действия, согласующегося с правилами игры.

3.2. От логических миров к гранулярным логическим матрицам

3.2.1. Логические миры и пространства

Понятие «логического пространства» было введено Л. Витгенштейном в его знаменитом «Логико-философском трактате» [5]. Автом за основу берется словосочетание «возможное положение дел» (Possible State of Affairs), которое описывается некоторым высказыванием. С каждым элементарным положением дел соотносится ось логиче-

ского пространства. Точкам логического пространства соответствуют некоторые состояния (полные состояния по фон Вригту [3] или «возможные миры» [26]). Любому предложению сопоставляется некоторая область логического пространства, т.е. множество возможных миров.

Те положения дел, которые имеют место в действительности, называются фактами; им соответствуют истинные высказывания. Тогда логический мир определяется как совокупность фактов. Условие логической независимости положения дел означает, что в мире можно получить любые их комбинации. Если число положений дел конечно и равно n , то общее число таких комбинаций равно 2^n . Любую такую комбинацию можно назвать «возможным миром». В целом, положения дел характеризуют конфигурации предметов, а высказывания – комбинации имен.

В качестве образующих логического пространства могут выступать типы атомарных логических предложений (см., например, трехмерное логическое пространство Н.А. Васильева [1]), типы рассуждений, соответствующие различным режимам мышления (пространство А.С. Есенина-Вольпина [4]), и пр.

Само по себе понятие логического мира как множества логических объектов никак заранее не предопределяет количество этих объектов или количество элементов, образующих логический универсум. Иными словами, логическим миром может быть признана любая непустая совокупность истинностных значений [7, 14].

Элементы такой обобщенной совокупности могут удовлетворять тем или иным требованиям качественного характера, существенным для структурирования логического универсума в целом. Согласно Я.В. Шрамко [14], среди этих возможных требований два представляются наиболее важными: а) принцип различия (объекты логического мира должны отличаться друг от друга); б) принцип выделенности (некоторые из них должны иметь особый статус, т.е. быть выделенными).

Тогда для представления простых логических миров вполне достаточно указать состав и мощность множеств значений истинности V и выделенных значений истинности D , где $D \subset V$. Ниже будем представлять логический мир как универсум истинностных значений V_i вместе с множеством выделенных значений D_i , понимаемых как абстрактные логические объекты, т.е. парой $\langle V_i, D_j \rangle$, $i=1,2,\dots,n,\dots,\infty$, $j < i$. Так классический логический мир Г. Фреге задается в виде $\langle V_2 = \{T, F\}, D_1 = \{T\} \rangle$, а логический мир модализированных истинностей Н. Решера определяется как $\langle V_4 = M_2 \times V_2, D_1 \rangle$, где $M_2 = \{L, Q\}$, L – «необходимо», Q – «случайно» (или иначе $M_2^* = \{P, Z\}$, P – «разрешено», Z – «запрещено»).

Логический мир Клини-Лукасевича задается парой $\langle V_3 = \{T, N, F\}, D_1 = \{T\} \rangle$ где N – «ни истина, ни ложь» (например, неопределенность у Клини и возможность у Лукасевича), а логический мир Васильева-Приста – в виде $\langle V_3 = \{T, B, F\}, D_2 = \{T, B\} \rangle$, где B – «и истина, и ложь одновременно». Наконец, логический мир Данна-Белнапа [16,18] определяется как $\langle V_4 = \{T, B, N, F\}, D_2 \rangle$.

Пространством называется множество, имеющее структуру, которая задается аксиоматически. Иными словами, пространство есть пара $\langle X, r \rangle$, где X – множество объектов, а r – некоторое отношение между ними. Например, метрическое пространство записывается как $\langle X, \rho \rangle$, где отношение ρ есть метрика, удовлетворяющая аксиомам антирефлексивности, симметричности и транзитивности, а пространство приближений есть пара $\langle X, \varphi \rangle$, где φ – отношение неразличимости (эквивалентности), которое является рефлексивным, симметричным и транзитивным.

В соответствии с подходом Данна-Шрамко [14, 18] будем определять обычное логическое пространство как пару:

$$LS = \langle V, 2^V \rangle, \quad (2)$$

где фиксированное множество значений истинности V рассматривается вместе с множеством своих подмножеств 2^V . Иными словами, множества истинностных значений будем называть логическим пространством, если на нем выделены определенные подмножества.

Естественными обобщениями формулы (2) являются: многозначное (n -значное) логическое пространство $\langle V, 2^V \rangle$, где n – целое число, $n > 2$; нечеткое логическое пространство Заде $\langle [0,1], [0,1]^V \rangle$, где $[0,1]^V$; есть множество нечетких подмножеств; L -нечеткое логическое пространство Гогена $\langle L, L^V \rangle$ и т.п. Будем использовать формулу (2) и ее обобщения для получения гранулярных значений истинности.

3.2.2. Логические матрицы

Важнейшим инструментом для семантического определения логических систем служат логические матрицы Лукасевича-Тарского, обозначаемые LM . Под *логической матрицей* понимается тройка:

$$LM = \langle V, O, D \rangle, \quad (3)$$

где V есть непустое множество значений истинности; O – множество операций над значениями истинности v из V ; $D \subset V$ – множество выделенных значений истинности. В общем случае множество D выражается в виде: $D = D^+ \cup D^-$, $D^+ \cap D^- = \emptyset$, где D^+ – множество собственно выделенных значений (значений, «подобных истине», а D^- – множество антивыведенных значений (значений, «подобных лжи»).

Очевидно, что логическая матрица (3) может быть представлена в виде пары:

$$LM = \langle UA, D \rangle, \quad (3^*)$$

где UA – универсальная алгебра с сигнатурой Ω . Иными словами, логическая матрица задает алгебраическое представление логики.

Отметим, что в логической матрице Тарского предполагается возможность получения истинностной оценки высказывания вне зависимости от контекста, от условий оценки. В качестве примеров можно привести алгебраическое представление классической логики $LM_{C2} = \langle \{0, 1\}, \{\neg, \wedge, \vee, \rightarrow\}, \{1\} \rangle$, трехзначной логики Клини $LM_{K3} = \langle \{0, 0.5, 1\}, \{\neg, \wedge, \vee, \rightarrow_K\}, \{1\} \rangle$, где промежуточное значение истинности 0.5 понимается как «неопределенность», а основные логические операции (отрицание, конъюнкция, дизъюнкция, импликация соответственно) задаются с помощью таблиц истинности, и бесконечнозначной логики Лукасевича $LM_{L\infty} = \langle [0, 1], \{\neg, \wedge, \vee, \rightarrow_L\}, \{1\} \rangle$, где промежуточные значения есть градации возможности, и $\forall v, w \in [0, 1], \neg v = 1 - v, v \wedge w = \min(v, w), v \vee w = \max(v, w), v \rightarrow_L w = \min(1, 1 - v + w)$.

В [8] рассмотрены такие пути расширения классической логической матрицы как: 1) введение обобщенных значений истинности, например, с помощью решеток и бирешеток; допущение интервальных, нечетких и лингвистических значений истинности, в том числе, интервалов и нечетких интервалов выделенных значений; 2) введение дополнительных операций над значениями истинности; нестандартное функционально-аксиоматическое задание этих операций.

Вариантом простого расширения логической матрицы для классической логики LM_C является логическая матрица для модальной логики $LM_{ML} = \langle \{0, 1\}, \{\neg, \wedge, \vee, \rightarrow, \Box\}, \{1\} \rangle$, где $\Box$ – сильный положи-

тельный модальный оператор. Примером переосмысления LM_{C2} служит логическая матрица для нечеткой логики, основанной на треугольных нормах T и треугольных конормах S [17, 24]:

$$LM_{GFL} = \langle [0, 1], \{n, T, S, I\}, [\alpha, 1] \rangle, \quad (4)$$

где n – произвольная операция отрицания, например $n(v) = 1 - v$, $\forall v \in [0, 1]$, $I(v, w) = S(n(v), w)$, $\forall v, w \in [0, 1]$, а α – некоторый порог истинности, $\alpha \in (0.5, 1)$.

Среди расширений логических матриц особый интерес представляют недетерминированные логические матрицы А. Аврона [15, 27].

Развитие аппарата логических матриц в плане построения гранулярных логических структур может опираться на обобщенную матрицу Вуйцицкого GLM_W :

$$GLM_W = \langle A, C \rangle, \quad (5)$$

где A – алгебра соответствующего типа с универсумом V , а C – произвольное семейство подмножеств A .

Таким образом, общий подход к построению и исследованию гранулярных семантик связан с формированием гранулярных матриц вида:

$$GLM_\Gamma = \langle A, \Gamma^V \rangle, \quad (5^*)$$

где Γ^V есть семейство подмножеств множества V , образующее базис семантической грануляции (например, $\Gamma^V = 2^V$ – множество четких подмножеств значений истинности или $\Gamma^V = [0, 1]^V$ – множество нечетких подмножеств значений истинности). В частности, удобно представить матрицу GLM_Γ в виде:

$$GLM_L = \langle L, \Gamma_L \rangle, \quad (5^{**})$$

где L – решетка, а Γ_L – гранулярная структура, получаемая как семейство подрешеток решетки L .

В теории решеток важную роль играют понятия фильтра и главного фильтра. Так, например, выделенные значения в многозначных

логики образуют фильтр по отношению к порядку значений истинности. Для строгого определения логической бирешетки опишем понятие бифильтра. Бифильтр F бирешетки BL есть непустое подмножество $F \subset BL$, такое, что $x \wedge y \in F \Leftrightarrow x \in F$ и $y \in F$, $x \otimes y \in F \Leftrightarrow x \in F$ и $y \in F$. Если к тому же выполняются условия $x \vee y \in F \Leftrightarrow x \in F$ или $y \in F$, $x \oplus y \in F \Leftrightarrow x \in F$ или $y \in F$, то бифильтр называется простым.

Логическая бирешетка [15] есть пара:

$$LBL = \langle BL, B\Phi \rangle \quad (6)$$

где BL – бирешетка, а $B\Phi$ – простой бифильтр, определенный на BL .

Таким образом, логические бирешетки (7) – это вариант логических матриц, где выделенные значения истинности описываются бифильтрами.

Алгебраические представления логических систем, изложенные выше, по сути, сохраняют вид логической матрицы Лукасевича-Тарского, расширяя те или иные ее компоненты: само понятие истинности, множество значений истинности, понятие выделенного значения, операций над значениями истинности.

В этой связи несомненный интерес представляет переосмысление подхода Тарского, связанное с учетом релятивизма логики, зависимости ее вида от онтологических предпосылок. Подобно тому, как семиотическая система Д.А. Поспелова была введена с целью обобщения понятия формальной системы на случай открытых и динамических сред, логико-семиотическая матрица [8], опирающаяся на идею логического пространства Л. Витгенштейна, а, в более общем плане, на концепцию семиотического пространства Ю.М. Лотмана, призвана показать контекст построения и область применимости описываемой логики. Здесь большие перспективы открывает изучение геометрии логических и семиотических пространств в контексте Эрлангенской программы Ф. Клейна.

Список литературы

1. *Васильев Н.А.* Воображаемая логика. М.: Наука, 1989.
2. *Витгенштейн Л.* Логико-философский трактат: Пер. с нем. М.: Наука, 1958.
3. *Вригт фон Г.Х.* Логико-философские исследования. Избранные труды: Пер. с англ. М.: Прогресс, 1986.

4. *Есенин-Вольпин А.С.* Избранное. М.: РГГУ, 1999.
5. *Заде Л.* Понятие лингвистической переменной и его применение к принятию приближенных решений: Пер. с англ. М.: Мир, 1976.
6. *Карнап Р.* Значение и необходимость. Исследование по семантике и модальной логике. М.: ЛКИ, 2007.
7. *Карпенко А.С.* Развитие многозначной логики. Изд. 3-е перераб. и доп. М.: Изд-во ЛКИ, 2010.
8. *Тарасов В.Б.* От алгебраической модели Тарского к логико-семиотической матрице // Труды 9-й национальной конференции по искусственному интеллекту КИИ-2004 (Тверь, 28 сентября – 2 октября 2004 г.). М.: Физматлит, 2004. Т. 1. С. 311–322.
9. *Тарасов В.Б.* О путях построения нестандартных и гибридных нечетких логик // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник трудов IV-й Международной научно-практической конференции (Коломна, 28–30 мая 2007 г.). М.: Физматлит, 2007. Т. 1. С. 101–113.
10. *Тарасов В.Б.* Грануляция информации, нестандартные и гибридные нечеткие множества // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник трудов VI-й Международной научно-практической конференции (Коломна, 16–19 мая 2011 г.). М.: Физматлит, 2011. Т. 1. С. 35–49.
11. *Тарасов В.Б.* От логических к диалогическим бирешеткам: использование в теории агентов // Вестник Ростовского государственного университета путей сообщения. 2011. № 3. С. 127–139.
12. *Фреге Г.* Логика и логическая семантика / пер. с нем. Изд. 2-е, испр. М.: Книжный дом «ЛИБРОКОМ», 2012.
13. *Шрамко Я.В.* Обобщенные истинностные значения: решетки и мультирешетки // Логические исследования. Вып. 9. М.: Наука, 2002. С. 264–291.
14. *Шрамко Я.В.* Истина и ложь: что такое истинностные значения и для чего они нужны? М.: Логос, 2009. № 2(70). С. 96–121.
15. *Arieli O., Avron A.* Reasoning with Logical Bilattices // Journal of Logic, Language and Information. 1996. Vol. 5. № 1. Pp. 25–63.
16. *Belnap N.* A Useful Four-Valued Logic // Modern Uses of Multiple-Valued Logic / Ed. by J.M. Dunn and G. Epstein. Dordrecht: D. Reidel, 1977. Pp. 8–37.
17. *Cignoli R., Esteva F., Godo L., Torrens A.* Basic Fuzzy Logic Is the Logic of Continuous T-Norms and Their Residua // Soft Computing. 2000. Vol. 4. Pp. 106–112.

18. *Dunn J.M.* Intuitive Semantics for First-Degree Entailment and Coupled Trees // *Philosophical Studies*. 1976. Vol. 29. Pp. 149–168.
19. *Dunn J.M.* Partiality and its Dual // *Studia Logica*. 2000. Vol. 66. Pp. 225–256.
20. *Fitting M.* Bilattices and the Theory of Truth // *Journal of Philosophical Logic*. 1989. Vol. 19. Pp. 225–256.
21. *Ginsberg M.* Multivalued Logics: a Unified Approach to Reasoning in AI // *Computer Intelligence*. 1988. Vol. 4. Pp. 256–316.
22. *Goguen J.* L-Fuzzy Sets // *Journal of Mathematical Analysis and Applications*. 1967. Vol. 18. Pp. 145–174.
23. *Handbook of Granular Computing* / Ed. by W. Pedrycz, A. Skowron, V. Kreinovich. New York: Wiley InterScience, 2008.
24. *Hajek P.* *Metamathematics of Fuzzy Logic*. Dordrecht: Kluwer Academic Publishers, 1998.
25. *Hintikka J.* *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*. Oxford: Clarendon and Oxford University Press, 1973.
26. *Kripke S.* Outline of a Theory of Truth // *Journal of Philosophy*. 1975. Vol. 72. Pp. 690–716.
27. *Logica Universalis. Towards a General Theory of Logic*/Ed. by J.-Y. Beziau. 2-nd ed. Basel: Birkhäuser Verlag, 2007.
28. *Lorenzen P., Lorenz K.* *Dialogische Logik*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1978.
29. *Lukasiewicz J.* *Selected Works* / Ed. by L. Borkowski. Amsterdam: North-Holland, 1970.
30. *Rescher N.* *Many-Valued Logic*. New York: McGraw-Hill, 1969.
31. *Svyatkina M.N., Tarassov V.B.* On Granules and Granular Structures for Multi-Valued Logics // *Handbook of the 4th World Congress and School on Universal Logics (UNILOG'2013, Rio de Janeiro, Brazil, March 29-April 07, 2013)* / Ed. by J.-Y. Beziau et al. Rio de Janeiro: Editora Kelps, Goiania, 2013. Pp. 297–300.
32. *Universal Logic: An Anthology: From Paul Hertz to Dov Gabbay*/Ed. by J.-Y. Beziau. Basel: Birkhäuser Verlag, 2007.
33. *Zadeh L.A.* Fuzzy Sets // *Information and Control*. 1965. Vol. 8. Pp. 338–353.
34. *Zadeh L.A.* Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic // *Fuzzy Sets and Systems*. 1997. Vol. 90. Pp. 111–127.

ГЛАВА 4

МОДЕЛИРОВАНИЕ МНЕНИЙ И ОЦЕНОК ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ: ОТ ЧЕТЫРЕХЗНАЧНЫХ СИСТЕМ МОДАЛЬНОСТЕЙ К НЕКЛАССИЧЕСКИМ МЕРАМ И НЕЧЕТКИМ РАСПРЕДЕЛЕНИЯМ

В.Б. Тарасов

Введение

В последние годы наблюдается становление системно-синергетического подхода в искусственном интеллекте (ИИ) [1–4]. С одной стороны, в русле развития теории интеллектуальных агентов, все большее значение имеет моделирование интенциональных характеристик агентов, в частности, желаний, намерений, обязательств и, прежде всего, мнений. С другой стороны, речь идет о возрождении давних традиций диалогизма в логиках для ИИ [5], обусловленном развитием систем диспута, аргументации, компьютерной диалектики, поддержки переговоров и т.п. В отличие от стандартной вертикальной логической структуры «понятие – суждение – рассуждение», соответствующей монологу одинокого агента, здесь на первый план выходят горизонтальные структуры, связанные с формированием и обменом мнениями в процессе диалога ради достижения истины или заключения соглашений.

Таким образом, если обычная логика (монологика) определяется как: наука о человеческом (нормативном) мышлении, его законах и формах; наука о правильных рассуждениях или (в семантическом аспекте) как наука о наиболее общих законах существования истины, то логика диалога (диалогика) обращает главное внимание не столько на извлечение заключений из посылок, сколько на увеличение (или, напротив, уменьшение) уверенности в этом заключении благодаря сопоставлению и согласованию разных мнений (с помощью утверждений и возражений, вопросов и ответов, гипотез и сомнений, аргументов и контраргументов).

Соответственно, моделирование понятия «мнение» (Belief) и тесно связанных с ним понятий суждений, аргументов, оценок, свидетельств, норм становится одной из центральных проблем ИИ. В общем случае, мнение автономного агента, которое может описывать состояния мира, других агентов и самого агента, обладает свойствами мягкости, гранулярности, градуированности; как правило, оно является частичным, неоднозначным, неточным, а подчас и противоречивым.

В 1975 г. Л. Заде в работе «Исчисление нечетких ограничений» [6] ввел понятие нечетких, приближенных высказываний, которые определяются путем наложения *нечетких ограничений* на значения исходных переменных. Позже он опубликовал работу «Нечеткие множества и информационная гранулярность» [7], в которой были описаны *информационные гранулы* как предложения вида $(X \text{ is } G) \text{ is } \alpha$, где G – нечеткое ограничение (гранула, т.е. группа объектов, связанных между собой отношением неразличимости, подобия, близости), а α – оценка возможности (или правдоподобия). Им было также отмечено, что значения данных внутри гранулы как бы склеиваются и становятся неотделимыми друг от друга.

Понятия *информационных гранул* и *обобщенных ограничений* (где наряду с нечеткими множествами, понимаемыми как функции распределения возможностей, могут использоваться распределения вероятности, значения истинности, модальности, нечеткие графики и т.д.) лежат в основе недавно предложенной *общей теории неопределенности* Л. Заде [8]. В дополнение к трем известным теориям информации А.Н. Колмогорова – комбинаторной, вероятностной, алгоритмической – в ней вводится четвертая концепция информации, связанная с распространением обобщенных ограничений.

Пятнадцать лет назад Л. Заде ввел концепцию *мягких вычислений* [9], объединяющих подходы нечеткой логики, нейроинтеллекта и эволюционного моделирования. При этом обычно нечеткие продукционные модели реализуются в обучаемых нейронных сетях, оптимизация которых происходит с помощью генетических алгоритмов. Затем в [10] А.Н. Аверкин и С.В. Прокопчина описали понятие *мягких измерений* (в этом контексте ими рассматривались, в частности, пути объединения вероятностных и нечетких измерений). В свою очередь, В.Б. Тарасов в статьях [4, 11] предложил понятие и пути формализации *мягких оценок*, которые тесно связаны с формированием мнений.

В настоящей работе обсуждаются варианты описания мнений с помощью гранулярных значений истинности, разных типов модаль-

ностей, неклассических мер. Изложение ведется в русле определения взаимосвязей между интенциональными, когнитивными, коммуникативными процессами для автономных интеллектуальных агентов и многоагентных систем [3].

4.1. Мнения и модальности

4.1.1. Мнения

Проблема исследования категории «мнения» (Belief) имеет давние и глубокие традиции. Еще Платон ясно осознавал наличие промежуточных градаций, ступеней формирования знания. Он различал собственно знание и мнение: «мнение не есть знание, но и не есть незнание: оно темнее знания и светлее незнания» [12]. Соответственно, в [13, 14] подчеркнуто, что знаниями в классическом ИИ называются «обоснованные, истинные убеждения» (мнения).

Для того чтобы суждение или умозаключение стало знанием, оно должно пройти проверку на полноту и непротиворечивость. Однако, реальные человеческие высказывания, оценки или рассуждения «здорового смысла» редко удовлетворяют этим жестким требованиям, определяющим статус знаний в ИИ. Здесь речь идет скорее о мнениях (точках зрения), которые отражают локальные представления о мире, пронизанные НЕ-факторами [4, 15]. Мнения носят не достоверный, а правдоподобный характер, зависят от ситуации и нередко могут пересматриваться.

Мнение может иметь дескриптивную и нормативную (модальную, ценностную) составляющие. Его описание возможно в системе *доксастических модальностей*: «уверен», «предполагаю», «сомневаюсь», «отвергаю». В частности, исходное мнение агента обычно выступает в форме *предположения* (гипотезы) – возможного, предварительного представления агента о мире и других агентах, которое выдвигается в интересах запуска механизма действия, направленного на его подтверждение, т.е. формирование *уверенности*.

Взаимодействие, столкновение разных мнений является основой споров (диспутов). В ходе спора агенты приводят аргументы в поддержку своего мнения и критикуют мнение противника. Различаются споры о ценностях (цель – достижение победы) и споры об истине (цель – установление истины).

В теории агентов различаются индивидуальные и коллективные мнения. Обычный жесткий язык описания системы мнений средствами дедуктивной логики состоит из аксиом, правил вывода и стратегии управления [16]. По сути, аксиомы отражают факты, в которых агент уверен (достоверные суждения). Напротив, правила вывода позволяют связать несколько мнений (в том числе, мнения о разных вещах). Стратегия управления квалифицирует мнение как истинное или ложное в зависимости от аксиом и правил вывода.

Таким образом, стандартный язык логики мнений представляет собой расширение языка классической логики с помощью модального оператора «Уверен» (Bel). Формула $Bel(a, p)$ читается «агент a уверен, что p », где суждение $p \in \mathcal{P}$. Формализация опирается на следующие положения.

1. Мнения агента являются согласованными, т.е. не могут одновременно рассматриваться абсолютно противоположные суждения p и $\neg p$.

2. Если имеем $Bel(a, p_1), Bel(a, p_2), \dots, Bel(a, p_n)$, и

если $p_1 \wedge p_2 \wedge \dots \wedge p_n$, то p ,

то $Bel(a, p)$

3. Правила могут быть неполными

Коллективные мнения агентов в многоагентной системе (МАС) могут быть явными и неявными, распределенными и разделенными, навязанными и общими [17]. Например, *общее мнение* – это мнение, которое принимается (разделяется) всеми агентами МАС, т.е. все агенты a_i из множества A уверены, что выполняется формула p : $Bel_C(A, p) = Bel(a_1, p) \wedge Bel(a_2, p) \wedge \dots \wedge Bel(a_n, p)$. Напротив, *неявное* (имплицитное) *мнение* обычно навязывается одним агентом всем другим агентам, причем справедлива формула: $Bel_I(A, p) = Bel(a_1, p) \vee Bel(a_2, p) \vee \dots \vee Bel(a_n, p)$. В свою очередь, *распределенное мнение*, соответствующее парадигме распределенного искусственного интеллекта (см., например, [3]), означает, что формула p является разложимой на подформулы p_i , $i = 1, \dots, n$, причем каждый агент a_i из A уверен, что выполняется p_i : $Bel_D(A, p) = \exists Bel(A, p) \mid \text{если } p_1 \wedge p_2 \wedge \dots \wedge p_n, \text{ то } p$.

Следует отметить, что жесткий логический подход не учитывает многих нюансов формирования и использования мнений, в частности, их градуированности, неточности, частичной противоречивости и пр.

В общем случае, будем описывать мнения на основе градуированных истинностных суждений вида $v(p)$, $v \in V$, например, $v \in [0, 1]$,

или близких к ним по форме градуированных модальных оценок $m(p)$. Так мнение эксперта «весьма возможно, что осенью доллар очень сильно упадет по отношению к евро» можно представить в виде *мягкой лингвистической оценки*:

$$km(X \text{ есть } cC),$$

где c – модификатор, позволяющий образовать составное лингвистическое значение ограничения на X , например, «очень» в составном терме «очень сильно», а k – аналогичный модификатор для некоторой модальности, например, квалификатор «весьма» в словосочетании «весьма возможно».

Ниже рассмотрим различные аспекты мнений в деятельности агентов, а также остановимся на методах представления мнений с помощью качественных и количественных (градуированных) модальностей.

4.1.2. Описания и предписания в деятельности агента

Функционирование любого интеллектуального агента опирается как на *описания*, так и на *предписания*. И описания, и предписания отражают информационные отношения между агентом и его средой: в первом случае это – информация о состояниях среды, воспринимаемых или представляемых агентом, а во втором случае – о возможных действиях агента на эту среду (задаваемая в форме приказов, команд, требований и пр.).

С одной стороны, представления агента о мире выражаются в виде *истинностных суждений* описательного характера, т.е. высказываний вида $p = X \text{ is } C$, характеризующихся значениями истинности $v \in V$. Обычно суждение является приближенным и правдоподобным, т.е. описывается в виде $v(p) = v(X \text{ есть } C)$, где C – гранулярное ограничение на X (интервал, распределение, словесная оценка). Суждение называется достоверным, если $v(p) = T(p)$. Помимо утверждений также могут использоваться отрицания $p = X \text{ is not } C$, противоречивые высказывания (по Н.А. Васильеву) $p = X \text{ is } C \text{ and is not } C$ и неопределенные высказывания $p = X \text{ is not } C \text{ and is not } (not C)$.

С другой стороны, любое *требование* или *предписание* агенту записывается аналогично в виде $r = a \text{ does } d$, $d \in D$ (агент a выполняет действие d из множества действий D). При этом обычно имеем мо-

дальнюю составляющую предписания r , $m(r)$, $m \in M$, например, $m_1(r)$ – «агент a должен выполнить действие d » или $m_2(r)$ – «агент a может выполнить действие d ».

Соотношение между описаниями и предписаниями показано на рис. 1. Из него видно, что истинность суждения может пониматься как соответствие между объектом внешнего мира и его описанием (первичен объект – дескриптивная модель). В то же время модальная оценка (ценность, полезность) предписания предполагает первичность какого-либо требования, приказа или нормы, распространяемой на некоторый класс объектов (или агентов), т.е. связана с построением нормативной модели

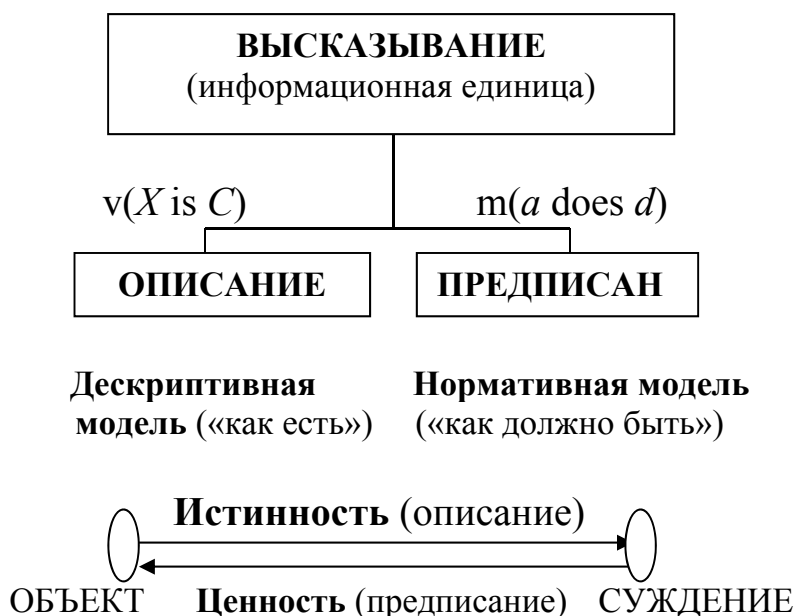


Рис. 1. Нормативные и дескриптивные логические модели

У агентов прагматические суждения опираются на стандарты, образцы, эталоны и т.п. При этом образец принципиально отличается от примера. *Пример* говорит о том, что имеет место в действительности, а *образец* – о том, что должно быть. Иными словами, примеры используются для поддержки описательных суждений, а ссылки на образцы служат обоснованием команд, предписаний и требований.

4.1.3. Ценности, оценки и нормы

В теории агентов центральное место занимает формализация предписаний, оценок, норм. По определению любой агент является *деятелем*, т.е. выполняет некоторые действия, которые соответству-

ют его роли в многоагентной системе, а также взаимосвязям между ролями. Любая роль характеризуется набором *обязательных* и *разрешенных* действий, которые могут инициироваться как собственными намерениями агента, так и *требованиями* со стороны других агентов. Статус агента в МАС определяется *оценкой результатов* его действий, причем эти действия должны не противоречить *нормам* поведения агентов в МАС

Оценкой (оценочным суждением) называется высказывание, которое устанавливает абсолютную или относительную (сравнительную) ценность некоторых объектов. Оценки выражаются с помощью аксиологических модальностей: абсолютных («отлично», «плохо», «хорошо», «удовлетворительно») или относительных («лучше», «хуже», «равноценно») [18, 19]. Любая оценка может быть задана в виде фрейма-прототипа:

{Оценка, ⟨Кто (агент оценки),...⟩, ⟨Что (Кого) (объект оценки),...⟩,
⟨Какая (тип оценки),...⟩, ⟨Почему (основание оценки),...⟩}.

Нормы – это социальные запреты и ограничения, накладываемые сообществом (многоагентной системой) на отдельного агента. Нормативное высказывание служит для установления некоторой нормы поведения. С одной стороны, нормы есть частный случай оценок: их можно рассматривать как общественно апробированные и закреплённые оценки. Средством, превращающим оценку в норму, является угроза наказания, т.е. стандартизация норм осуществляется с помощью санкций. Так еще К. Менгер установил прямую связь между предписанием и санкцией: «обязательно r , а если не r , то наказание или ухудшение». С другой стороны, формирование норм предполагает согласование разных мнений по этим нормам

Итак, реализация агентом нормативного поведения предполагает наличие, по крайней мере, двух элементов: *нормы*, обязательной для выполнения в данной ситуации, и *оценки* степени выполнения ее предписаний.

Формально норму как предписание к действию можно выразить четверкой:

$$NORM = \langle A, act, M, W \rangle,$$

где A – множество агентов, которым адресована норма;

$act \in ACT$ – действие, являющееся объектом нормативной регуляции (содержание нормы);

W – множество миров, в которых применима норма (условия приложения, обстоятельства, в которых должно или не должно выполняться действие);

M – множество систем модальностей, связанных с действием, например, система $M_N = \{O, P, B, Z\}$, O – «обязательно», P – «разрешено», B – «безразлично», т.е. «необязательно», Z – «запрещено».

4.1.4. Единый подход к описанию различных систем модальностей на основе четырехзначных семантик

При формализации агентов могут использоваться различные системы модальностей: доксистические модальности (модальности мнения), деонтические модальности (модальности действий и норм), оптитативные модальности (модальности желаний), эротетические модальности (модальности вопросов и ответов), оценочные модальности и пр. В связи с этим актуально построение единого подхода к формализации различных систем модальностей. Здесь подобный подход развивается на основе многозначных логических семантик.

Следуя А.А. Ивину, возьмем в качестве двух базовых характеристик любой модальности ее «знак» и «силу». Тогда получаем двухкоординатное представление систем модальностей (рис. 2), которое в простейшем случае сводится к четырехзначному множеству $M_4 = \{СП, СлП, СлО, СО\}$, где СП – сильная положительная модальность; СлП – слабая положительная модальность, СлО – *слабая отрицательная модальность*, СО – сильная отрицательная модальность. Первым трем модальностям в модальной логике соответствуют унарные операторы $\Box$, $\Diamond$, ∇ . Соотношения между ними хорошо известны: $\Box r = \nabla \Diamond \neg r$, $\nabla r = \Box \neg r$, где r – предписание.

Возьмем двухкоординатное представление систем модальностей в интервале $[-1, +1]$ (рис. 2). Тогда вышеуказанные модальности можно описать в виде пар: СП = $(+1, +1)$, СО = $(+1, -1)$, СлП = $(-1, +1)$, СлО = $(-1, -1)$.

Независимо от того, к какой группе относятся модальности, они определяются друг через друга по общей схеме (табл. 1). В частности, $\neg СП(r) = СлО(r)$ (отрицание сильной положительной модальности есть слабая отрицательная модальность), $\neg СлО(r) = СП(r)$ (отрицание слабой отрицательной модальности есть сильная положительная модаль-

ность), $\neg \text{СлП}(r) = \text{СлО}(r)$ (отрицание слабой положительной модальности есть сильная отрицательная модальность), $\neg \text{СлО}(r) = \text{СлП}(r)$ (отрицание сильной отрицательной модальности есть слабая положительная модальность).

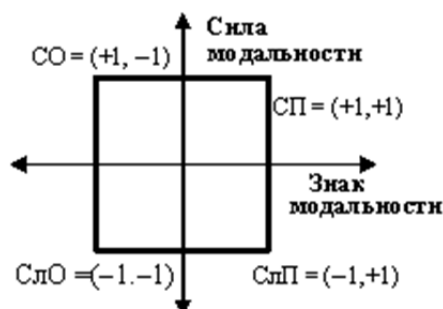


Рис. 2. Двухкоординатное представление модальностей

Таблица 1

Единый способ представления разных классов модальностей

Сильная положительная модальность (+1, +1)	Слабая положительная модальность (-1, +1)	Слабая отрицательная модальность (-1, -1)	Сильная отрицательная модальность (+1, -1)	Классы модальностей
Необходимость	Возможность	Случайность	Невозможность	Алетические
Уверенность	Предположение	Сомнение	Отвержение	Доксастические
Обязанность	Разрешение	Безразличие	Запрещение	Деонтические
Корректность	Противоречие	Некорректность	Нерелевантность	Эротетические

Примерами практической интерпретации этих определений являются связи между слабой нормой и сильной антинормой в деонтической логике: «то, что не разрешено, запрещено» и наоборот «то, что не запрещено, разрешено».

По другой общей схеме, связи между модальностями могут задаваться и с помощью бинарных операций, например, $\text{СлП}(r) = \text{СП}(r) \vee \text{СлО}(r)$ (слабая положительная модальность выражается как дизъюнкция сильной положительной и слабой отрицательной).

Для различных классов модальностей можно сформулировать принцип модальной непротиворечивости $\neg(\text{СП}(r) \wedge \text{СлО}(r)) = \neg \text{СП}(r) \vee \neg \text{СлО}(r)$.

Например, суждение не может быть логически необходимым, и логически невозможным; действие не может быть одновременно и обязательным, и запрещенным. Также выполняется соотношение $\neg(\Box(r) \vee \Box O(r)) = \neg\Box(r) \wedge \neg\Box O(r) = \Box\Pi(r)$, т.е. слабая отрицательная модальность есть антидизъюнкция сильной положительной и сильной отрицательной модальности, например, в логике мнений: «не уверен и не отвергает, значит сомневается»).

4.1.5. Решетки модальностей

Зададим на произвольном непустом конечном множестве модальностей M два различных отношения порядка $\leq_K$ и $\leq_U$ следующим образом: $\forall r, s \in M, r = (x_1, y_1), s = (x_2, y_2), r \leq_K s \Leftrightarrow x_1 \leq x_2 \text{ и } y_1 \leq y_2; r \leq_U s \Leftrightarrow x_1 \leq x_2 \text{ и } y_1 \geq y_2$. Интерпретации этих двух порядков могут быть различными, например, порядок сильной нормы и порядок слабой нормы, порядок необходимости (определенности) и порядок возможности, порядок уверенности и порядок предположения, порядок обязанности и порядок разрешения, и пр. Так для M_4 имеем $\Box O \leq_K \Box \Pi \leq_K \Box \Pi$, $\Box O \leq_K \Box O \leq_K \Box \Pi$ и $\Box O \leq_U \Box \Pi \leq_U \Box \Pi$, $\Box O \leq_U \Box O \leq_U \Box \Pi$.

Математическая структура $\langle M^2, \leq_K, \leq_U \rangle$ является *биупорядоченным множеством*, поскольку здесь $M^2 = \{-1, +1\}^2$ есть непустое множество, а $\leq_K$ и $\leq_U$ – два различных отношения порядка. Нетрудно убедиться в том, что в случае M_4 упорядоченные множества $\langle M_4, \leq_K \rangle$ и $\langle M_4, \leq_U \rangle$ образуют полные решетки $ML4$ и $ML4^*$, аналогичные логической решетке Белнапа и аппроксимационной решетке Скотта (здесь решетка $ML4^*$ получается из $ML4$ поворотом на 90°). Диаграммы Хассе для этих решеток изображены на рис.3.

Эти рассуждения легко расширить на случай решетки $[-1, +1]^2$.

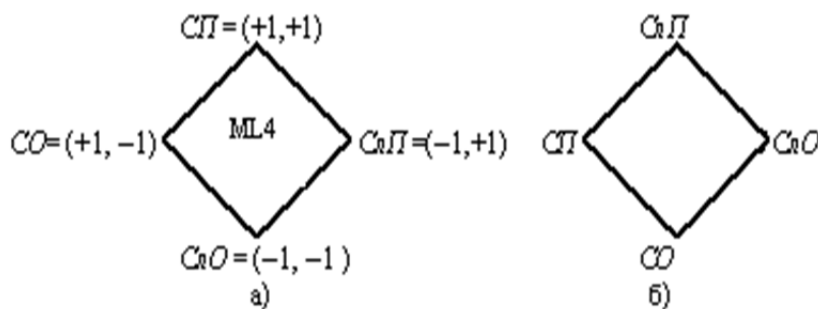


Рис. 3. Решетки модальностей: а) $ML4$; б) $ML4^*$

4.2. Неклассические меры и описание модальностей

Основными характеристиками любого множества являются *границы* и *мера*. Понятие *меры* есть одно из важнейших математических понятий, как, впрочем, и понятие интеграла, соответствующего данной мере. Оно является естественным обобщением понятия длины отрезка, площади плоской фигуры, объема пространственной фигуры. Классические меры удовлетворяют условию аддитивности.

Пусть A и B – некоторые события, а X – полное множество событий. *Мерой* называется функция множества:

$$m: 2^X \rightarrow \mathbb{R}^+, \quad \mathbb{R}^+ = [0, \infty),$$

которая удовлетворяет следующим условиям:

- (1) $\forall A \in 2^X, A \subseteq X \Rightarrow m(A) \geq 0$;
- (2) $m(\emptyset) = 0$;
- (3) $\forall A, B \in 2^X, m(A \cup B) = m(A) + m(B) - m(A \cap B)$.

4.2.1. Вероятностная мера

Наиболее известным случаем классической меры является *нормальная мера* или *вероятностная мера* А.Н. Колмогорова, которая выражается в виде:

$$P: 2^X \rightarrow [0, 1]$$

и удовлетворяет следующим условиям:

- 1) $P(\emptyset) = 0, P(X) = 1$ (ограниченность)
- 2) $\forall A, B \in 2^X, A \subseteq B \Rightarrow P(A) \leq P(B)$ (монотонность)
- 3) $\forall A, B \in 2^X, A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ (аддитивность).

В общем случае, берется σ -алгебра множеств в X , $\sigma \subseteq 2^X$ и аксиома аддитивности записывается в форме $\forall A_i \in \sigma, \bigcap A_i = \emptyset \Rightarrow P(\bigcup A_i) = \sum P(A_i)$. С вероятностной мерой связана статистика средних значений.

В теории меры хорошо известными примерами σ -аддитивных мер являются меры Лебега.

Пусть x_0 есть заданный элемент в X . Частным случаем вероятностной меры является примитивный класс мер Дирака m_D , определяемый соотношением: $\forall A \in 2^X$,

$$m_D(A) = \begin{cases} 1, & \text{если } x_0 \in A \\ 0 & \text{в противном случае.} \end{cases}$$

Мера Дирака есть частный случай вероятностной меры, соответствующий детерминированной сингулярной информации (мера полной уверенности).

4.2.2. Семейство мер Сугено [20, 21]

Требование аддитивности является слишком ограничительным для многих практических процедур оценивания и формирования мнений. Так еще в нашей работе [22]) была выдвинута гипотеза о том, что неаддитивность есть одно из фундаментальных отличий процедур оценивания от процедур измерения. При этом в качестве базы для оценивания предлагается пространство с *предмерой* (X, σ, u) , где *предмера* u удовлетворяет лишь условиям 1) и 2).

Тогда произвольная *псевдомера*, называемая также *неклассической (неаддитивной) мерой*, строится как однопараметрическое расширение обычной меры путем замены стандартной аксиомы 3) каким-либо более гибким условием, например, параметризованным λ -правилом Сугено ($u = g$) [20]:

$$(3^\circ) \forall A_i \in \sigma, g_\lambda(\cup A_i) = \Sigma g(A_i) + \lambda \Pi g(A_i), -1 \leq \lambda < \infty.$$

Это правило получается из уравнения $\lambda + 1 = \Pi(1 + \lambda^i)$.

Кроме условий 1), 2) и 3)', для меры Сугено выполняется также условие 4) $\forall A_n \in 2^X$, если $A_1 \subseteq A_2 \subseteq \dots$, то $\lim_{n \rightarrow \infty} g(A_n) = g(\lim_{n \rightarrow \infty} A_n)$.

Здесь при $\lambda > 0$ получаем семейство субаддитивных мер: $\forall A, B \in 2^X$, $g_\lambda(A \cup B) < g_\lambda(A) + g_\lambda(B)$, а при $-1 \leq \lambda < 0$ – семейство супераддитивных (синергетических) мер $g_\lambda(A \cup B) > g_\lambda(A) + g_\lambda(B)$. В случае $\lambda = 0$ мера Сугено превращается в обычную аддитивную (вероятностную) меру.

В общем случае, неклассические меры могут порождаться с помощью треугольных конорм S , моделирующих семантику связки «ИЛИ»

$$(3') \forall A_i \in \sigma, u(\cup A_i) = S(u(A_i)).$$

Такие меры принято называть S-разложимыми (по З. Веберу).

Двойственные им меры формируются на основе треугольных норм T [23].

4.2.3. Основы теории свидетельств

Одними из первых ученых, предложивших применять неклассические меры (псевдомеры) в интересах описания суждений (свидетельств) стали А. Демпстер и Дж. Шейфер [24]. Так Демпстер ввел функции верхних и нижних вероятностей, индуцируемых многозначными отображениями. В свою очередь, Шейфер построил теорию свидетельств на основе двух классов монотонных неаддитивных мер – мер доверия и мер правдоподобия (см. также [25]).

Мерой доверия называется монотонная функция множества:

$$b: 2^X \rightarrow [0,1],$$

удовлетворяющая следующим условиям:

- (a) $b(\emptyset) = 0, b(X) = 1$
- (б) $\forall A, B \in 2^X, b(A \cup B) \geq b(A) + b(B)$.

Здесь условие (б) определяет свойство *супераддитивности*.

Пусть A' есть дополнение A . Из определения меры доверия вытекает ее свойство $b(A) + b(A') \leq 1$ (субкомплементарность).

Если задана мера доверия, то двойственную к ней меру правдоподобия можно определить следующим образом:

$$Pl(A) = 1 - b(A'), \forall A \in 2^X \quad (1)$$

Монотонная мера правдоподобия Pl удовлетворяет следующим аксиомам:

- (a) $Pl(\emptyset) = 0, Pl(X) = 1$
- (б') $\forall A, B \in 2^X, Pl(A \cup B) \leq Pl(A) + Pl(B)$.

Аксиома (б') определяет условие *субаддитивности*.

Для меры Pl выполняется также условие суперкомплементарности $Pl(A) + Pl(A') \geq 1$.

Замечание 1. При переходе к нестандартным логическим семантикам Данна-Белнапа (т.е. заменяя 2^X на 2^V , где V – множество значений истинности), получаем интерпретацию меры истинности как функции доверия к высказыванию p , а меры ложности – как функции доверия к противоположному высказыванию $\neg p$.

Введем функцию $m_p: \wp \rightarrow [0,1]$, причем: 1) $m_p(\emptyset) = 0$; 2) $\sum_{p \in \wp} m_p(p) = 1$.

Тогда для любых $p, q \in \wp$ по Шейферу [24] получаем:

$$v(q) = b(q) = \sum_{p \text{ влечет за собой } q} m_p(p).$$

Аналогично с учетом условия (I) получаем:

$$Pl(q) = \sum_{p \text{ не влечет за собой } q} m_p(p)$$

Если рассматривать меры $b(A)$ и $Pl(A)$ в контексте теории количественных модальностей, то можно определить также меру недоверия $nb(A) = 1 - b(A)$ и меру отвержения (неправдоподобности) $nPl(A) = 1 - pl(A)$.

4.2.4. Меры возможности и необходимости

Из аксиомы монотонности для любой предмеры непосредственно вытекают два важных неравенства, характеризующие два фундаментальных класса псевдомер:

$$\begin{aligned} g(A \cup B) &\geq \max \{g(A), g(B)\} \\ g(A \cap B) &\leq \min \{g(A), g(B)\}. \end{aligned}$$

Тогда в граничных случаях определяются мера возможности Π как минимальная мера правдоподобия (предложена Л. Заде в [26]) и мера необходимости Дюбуа-Прада (см. [27]) как максимальная мера доверия.

Мера возможности есть функция множества:

$$\Pi: 2^X \rightarrow [0,1],$$

для которой справедливы условия:

1. $\Pi(\emptyset) = 0, \Pi(X) = 1$ (ограниченность)
2. $\forall A, B \in 2^X, A \subseteq B \Rightarrow \Pi(A) \leq \Pi(B)$ (монотонность)
3. $\forall A, B \in 2^X, \Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\}$ («либо-либо»-условие).

Меру Π можно задать на множестве высказываний $\wp$. Пусть $p, q \in \wp$. Тогда условие $\Pi(p \vee q) = \max\{\Pi(p), \Pi(q)\}$ можно интерпретировать следующим образом: истинность дизъюнкции двух суждений определяется возможностью появления хотя бы одного из них.

В свою очередь, нечеткое множество может пониматься как функция (плотность) распределения возможности:

$$\pi: X \rightarrow [0, 1]$$

удовлетворяющая условию нормировки:

$$\Pi(A) = \sup_{x \in A} \pi(x) = 1. \quad (2)$$

Аналогично мера необходимости есть функция множества:

$$N: 2^X \rightarrow [0, 1],$$

для которой выполняются требования:

1. $N(\emptyset) = 0, N(X) = 1$ (ограниченность)
2. $\forall A, B \in 2^X, A \subseteq B \Rightarrow N(A) \leq N(B)$ (монотонность)
- 3*. $\forall A, B \in 2^X, N(A \cap B) = \min\{N(A), N(B)\}$. («и-и» условие)

Если определить меру N на множестве высказываний $\wp$, то условие $N(p \wedge q) = \min\{N(p), N(q)\}$ означает, что истинность конъюнкции двух суждений определяется необходимостью их одновременного выполнения.

Для мер необходимости и возможности справедливо равенство (I):

$$N(A) = 1 - \Pi(A'), \quad \forall A \in 2^X$$

Условие (I) можно записать и в более общей форме:

$$N(A) = n(\Pi(A')), \quad (1^*)$$

где n – некоторая функция отрицания.

Из выражений (1) и (2) вытекает, что меру необходимости можно определить по функции распределения возможности:

$$N(A) = \inf_{x \notin A} (1 - \pi(x))$$

Соотношение между возможностью и необходимостью записывается в виде:

$$\Pi(A) \geq N(A)$$

В отличие от выполняемого для вероятностной меры закона $P(A) + P(A') = 1, \forall A \in 2^X$, для меры возможности имеем условие:

$$\Pi(A) + \Pi(A') \geq 1, \forall A \in 2^X,$$

а для меры необходимости:

$$N(A) + N(A') \leq 1, \forall A \in 2^X$$

Кроме того, из $\Pi(A) < 1$ следует $N(A) = 0$ (неполная возможность A приводит к абсолютной неуверенности), а из $N(A) > 0$ вытекает $\Pi(A) = 1$ (наличие некоторой уверенности в A означает абсолютную возможность).

В свою очередь, такие понятия как невозможность $n\Pi$ и проблематичность (не необходимость, случайность) nN легко описать с помощью обычного оператора отрицания на основе мер возможности и необходимости соответственно:

$$\begin{aligned} n\Pi(A) &= 1 - \Pi(A), \forall A \in 2^X \\ nN(A) &= 1 - N(A), \forall A \in 2^X \end{aligned}$$

Заметим, что на множестве высказываний $\wp$ $n\Pi(p \vee q) = \min\{n\Pi(p), n\Pi(q)\}$ (ложность понимается как невозможность дизъюнкции двух суждений, которая определяется невозможностью и одного, и другого), а $nN(p \wedge q) = \max\{nN(p), nN(q)\}$ (проблематичность конъюнк-

ции двух суждений определяется проблематичностью хотя бы одного из них).

В логике возможности слабый принцип модальной полноты можно выразить в виде $\max\{P(p), P(\neg p)\}=1, \forall p \in \wp$ (два противоположных суждения наверняка возможны) или $\max\{nP(p), nP(\neg p)\}=1, \forall p \in \wp$ (два противоположных суждения наверняка проблематичны). В свою очередь, слабый принцип модальной непротиворечивости можно представить в виде: $\min\{P(p), P(\neg p)\}=0, \forall p \in \wp$ (два противоположных суждения не могут быть необходимыми одновременно) или $\min\{nP(p), nP(\neg p)\}=0, \forall p \in \wp$ (два противоположных суждения являются невозможными одновременно)

Предельному случаю *полной определенности* события А соответствуют значения:

$$P(A)=1 \Leftrightarrow P(\neg A)=0 \text{ или } nP(A)=0 \Leftrightarrow nP(\neg A)=1,$$

а другому предельному случаю *полной неопределенности (незнания)* отвечают:

$$P(A) = P(\neg A) = 1 \Leftrightarrow P(A)=P(\neg A) = 0.$$

Пусть $A \in 2^X$ – произвольное событие, а $E \in 2^X$ – достоверное (всегда определенное) событие. Тогда в частном случае меры возможности, невозможности, необходимости, проблематичности могут определяться, исходя из анализа принадлежности элементов множества А множеству Е, что сводится к анализу следующих четырех вариантов [27]:

1) $A \cap E \neq \emptyset$; 2) $A \cap E = \emptyset$; 3) $\neg A \cap E = \emptyset$, т.е. $E \subseteq A$; 4) $\neg A \cap E \neq \emptyset$,

т.е. $E \not\subseteq A$

Тогда мера возможности события А со значениями в $\{0,1\}$ может быть задана как:

$$P_E(A) = \begin{cases} 1, & \text{если } A \cap E \neq \emptyset \\ 0, & \text{в противном случае.} \end{cases}$$

Мера необходимости может быть представлена в виде:

$$N_E(A) = \begin{cases} 1, & \text{если } E \subseteq A \\ 0, & \text{в противном случае.} \end{cases}$$

В частном случае, мера невозможности может быть представлена в виде:

$$n\Pi_E(A) = \begin{cases} 1, & \text{если } A \cap E = \emptyset \\ 0, & \text{в противном случае,} \end{cases}$$

а мера случайности (проблематичности):

$$nN_E(A) = \begin{cases} 1, & \text{если } E \not\subseteq A \\ 0, & \text{в противном случае} \end{cases}$$

Замечание 2. По аналогии с единым способом представления разных классов модальностей на основе четырехзначных семантик (см. таблицу 1) здесь предлагается расширить интерпретацию мер возможности и необходимости, рассмотрев единый способ описания градуальных модальностей мнения или оценки на основе четырех мер $N(p)$, $\Pi(p)$, $n\Pi(p)$, $nN(p)$.

Упорядочим градуальные модальности следующим образом: сильная градуальная положительная модальность (СГП), слабая градуальная положительная модальность (СлГП), слабая градуальная отрицательная модальность (СлГО), сильная градуальная отрицательная модальность (СГО). Тогда:

$$\begin{aligned} \text{СГП}(p \wedge q) &= \min\{\text{СГП}(p), \text{СГП}(q)\} \\ \text{СлГП}(p \vee q) &= \max\{\text{СлГП}(p), \text{СлГП}(q)\} \\ \text{СлГО}(p \wedge q) &= \max\{\text{СлГО}(p), \text{СлГО}(q)\} \\ \text{СГО}(p \vee q) &= \min\{\text{СГО}(p), \text{СГО}(q)\}. \end{aligned}$$

Таким образом, можно расширить определения сильных положительных модальностей, например, деонтической модальности «обяза-

тельства» $O(p \wedge q) = \min\{O(p), O(q)\}$, слабых положительных модальностей, например, оптитативной модальности «желательности» $Des(p \vee q) = \max\{Des(p), Des(q)\}$, слабых отрицательных модальностей, например, докастической модальности «сомнение» $Dbt(p \wedge q) = \max\{Dbt(p), Dbt(q)\}$ и пр.

4.2.5. Качественные оценки возможности и нечеткости

Идея построения сравнительных оценок возможности восходит к работам Д. Льюиса, который интерпретировал возможность как отношение сходства. Затем Дюбуа и Прад показали, что мера возможности индуцирует отношение $\geq_{\Pi}$ между событиями: $A \geq_{\Pi} B$ тогда и только тогда, когда $\Pi(A) \geq \Pi(B)$ [27]. Здесь $A \geq_{\Pi} B$ означает, что возможность события A , по крайней мере, не меньше возможности события B . Отношение $\geq_{\Pi}$ обладает следующими свойствами: а) $T \geq_{\Pi} F$, где T и F – истина и ложь соответственно; б) $A \geq_{\Pi} B$ или $A \leq_{\Pi} B$ (сравнимость); в) $A \geq_{\Pi} B, B \geq_{\Pi} C \Rightarrow A \geq_{\Pi} C$ (транзитивность); г) если $B \geq_{\Pi} C$, то для любого A имеем $A \cup B \geq_{\Pi} A \cup C$.

В свою очередь, Трильяс и Альсина обобщили идею сравнительных оценок для произвольных неклассических мер, введя (рефлексивное и транзитивное) отношение предпорядка $\geq_g$. Здесь $A \geq_g B$ означает, что множество A обладает неким свойством в степени, не меньшей, чем множество B .

Отношение предпорядка по включению множеств позволяет с единых позиций описать не только неклассические меры, определенные на 2^X , но и функции нечетких множеств, заданные на $[0,1]^X$.

4.3. Меры на нечетких множествах

Различные меры на нечетких множествах можно определить, вводя разные отношения порядка (или предпорядка) на интервале $[0,1]$. Здесь классическое отношение порядка (порядок вложенности нечетких множеств) задается в виде: $\mu \leq \nu \Leftrightarrow \mu(x) \leq \nu(x), \forall x \in X$.

Рассмотрим максимально нечеткое множество с функцией принадлежности $\mu(x) = 0.5$. Тогда новое отношение порядка $\leq'$, называемое «порядком заострения» [25, 28], можно задать следующим обра-

зом: $\mu \leq' v \Leftrightarrow \mu(x) \leq' v(x), \forall x \in X$, где $\mu(x) \leq' v(x)$ тогда и только тогда, когда $\mu(x) \leq v(x)$ при $\mu(x) \leq 0.5$ и $\mu(x) \geq v(x)$ при $\mu(x) \geq 0.5$.

Отношениям порядка $\leq$ и $\leq'$ ставятся в соответствие два класса мер – меры энергии и меры энтропии нечетких множеств соответственно [28, 29].

Пусть высказывание $p \in \wp$. Как известно, противоречие в классической логике записывается в форме $p \wedge \neg p$. В обобщенном виде его можно выразить формулой $p T n(p)$, где T – треугольная норма, соответствующая лингвистической связке «И», а n – унарная операция отрицания в функционально-аксиоматической форме.

Введем отношение предпорядка, индицируемое отрицанием n , т.е. рефлексивное и транзитивное отношение $\leq_n$ на $[0, 1]$:

$$p \leq_n q \Leftrightarrow p T n(p) \leq q T n(q)$$

и будем рассматривать предупорядоченное множество $[0, 1]_n$.

Для наибольшей треугольной нормы $T = \min$ имеем $p \leq_n q \Leftrightarrow \min\{p, n(p)\} \leq \min\{q, n(q)\}$.

В случае оппозиционной шкалы отношение $\leq_n$ на интервале $[-1, +1]$ задает меру гармонии.

4.3.1. Меры энергии (показатели силы принадлежности)

Пусть X – базовое множество, на котором определено нечеткое множество $\mu: X \rightarrow [0, 1]$, а $[0, 1]^X = \{\mu \mid \mu: X \rightarrow [0, 1]\}$ – множество всех нечетких подмножеств.

Обозначим через $\mathbb{R}^+$ множество всех неотрицательных действительных чисел $\mathbb{R}^+$. Мерой энергии нечеткого множества называется функция:

$$e: [0, 1]^X \rightarrow \mathbb{R}^+,$$

удовлетворяющая следующим аксиомам:

- e1) $e(\mu) = 0$ тогда и только тогда, когда $\mu(x) = 0$ для всех x из X ;
- e2) $e(\mu)$ принимает максимальное значение тогда и только тогда, когда $\mu(x) = 1$ для всех x из X ;
- e3) $\forall \mu, v \in [0, 1]^X, \mu(x) \leq v(x) \Rightarrow e(\mu) \leq e(v)$.

4.3.2. Меры энтропии

Пусть X – базовое множество, на котором определено нечеткое множество $\mu: X \rightarrow [0,1]$, а $[0,1]^X = \{\mu \mid \mu: X \rightarrow [0,1]\}$ – множество всех нечетких подмножеств. Мера энтропии [28] определяется в виде функции:

$$h: [0,1]^X \rightarrow \mathbb{R}^+,$$

удовлетворяющей следующим условиям:

h1) $h(\mu) = 0$ тогда и только тогда, когда $\mu(x) = f(x) \in \{0,1\}$, т.е. когда f – классическая характеристическая функция множества;

h2) $h(\mu) = h_{\max}$ тогда и только тогда, когда $\mu(x) = 0.5$ для всех $x \in X$;

h3) $\forall \mu, \nu \in [0,1]^X, \mu(x) \leq \nu(x) \Rightarrow h(\mu) \leq h(\nu)$.

Известны и другие определения энтропии, в частности, энтропии по А. Кофману, как нормализованного расстояния до предельно нечеткого распределения $\mu(x) = 0.5, \forall x \in X$, а также как расстояния между нечетким множеством и его дополнением [30]. Согласно И.З. Батыршину, мера энтропии на алгебре может пониматься как мера ее небулевости.

В общем случае энтропию можно определить через отношение предпорядка $\leq_n$ как функцию:

$$h(\mu) = k \sum_{x \in X} \{T(\mu(x), n(\mu(x)))\},$$

где T и S – треугольная норма и конорма соответственно, n – операция отрицания, а k – константа (коэффициент нормализации).

4.3.3. Меры специфичности

Меры специфичности (неспецифичности) нечетких множеств тесно связаны с понятием гранулярности и показывают степень точности задания нечеткого множества

Пусть X – базовое множество, а $[0,1]^X = \{\mu \mid \mu: X \rightarrow [0,1]\}$ – множество всех нечетких подмножеств, определенных на X .

Мера специфичности по Р. Ягеру [31] есть нормализованная функция нечеткого множества:

$$sp: [0,1]^X \rightarrow [0,1],$$

такая что:

sp1) $sp(\mu) = 1$ тогда и только тогда, когда μ есть одноточечное множество, $\mu = \{x_i\}$;

sp2) $sp(\mu) = 0$, если μ – пустое множество;

sp3) $\forall \mu, \nu \in [0, 1]^X, \mu(x) \leq \nu(x) \Rightarrow sp(\mu) \leq sp(\nu)$.

Обратная мера неспецифичности была введена Хигаши и Клиром. В свою очередь, Дюбуа и Прад предложили принцип минимальной специфичности как основу рассуждений на основе мер доверия и правдоподобия, дополняющий известный принцип максимальной энтропии.

Заключение

В работе рассмотрен широкий круг вопросов, связанный с вариантами описания интенциональных характеристик интеллектуальных агентов. Во-первых, проведено детальное обсуждение понятия «мнение» и его связей с оценками, нормами, свидетельствами. Отмечена необходимость учета как описаний, так предписаний при формализации мнений агентов, предложены варианты их моделирования. Во вторых, особое внимание уделено модальной трактовке мнения на базе системы доксистических модальностей. Предложен единый подход к интерпретации разных систем модальностей в русле многозначных и нечетких логик. На конечном множестве модальностей M введены два различных отношения порядка; построены решетки модальностей, определяемые этими отношениями. В третьих, осуществлен естественный переход от описания мнений с помощью обычных модальностей к их градуальному описанию на основе неклассических мер и соответствующих функций распределения. Проведен анализ основных неклассических мер, прослежены связи между мерами возможности и необходимости и одной стороны и показателями пересечения и включения множеств с другой. Рассмотрены способы построения нестандартных логических семантик с помощью мер доверия и правдоподобия. Также предложены варианты формализации сравнительных модальных оценок, опирающихся на качественные псевдомеры. В четвертых, изучены взаимосвязи между мерами, определенными на обычных и нечетких множествах.

В результате систематизирован существующий инструментарий и предложены новые подходы к моделированию интенциональных характеристик автономных и коллективных агентов.

Список литературы

1. *Тарасов В.Б.* Системно-организационный подход в искусственном интеллекте // Программные продукты и системы. 1997. № 3. С. 6–13.
2. *Тарасов В.Б.* Синергетический искусственный интеллект: истоки и перспективы // Труды Международного конгресса «Искусственный интеллект в XXI-м веке» (ICAI'2001, Дивноморск, Россия, 3–8 сентября 2001 г.). М.: Наука, Физматлит, 2001. С. 559–570.
3. *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям. М.: Эдиториал УРСС, 2002.
4. *Тарасов В.Б.* От нечетких множеств к мягким оценкам и синергетическому искусственному интеллекту // Сборник докладов Международной конференции по мягким вычислениям и измерениям (SCM'2005, Санкт-Петербург, СПбГЭТУ, 27–29 июня 2005 г.). СПб: СПбГЭТУ, 2005. С. 24–32.
5. *Тарасов В.Б.* От монологических к диалогическим подходам в искусственном интеллекте // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сборник трудов V-й Международной научно-практической конференции (Коломна, 28–30 мая 2009 г.). М.: Физматлит, 2009. Т. 1. С. 149–162.
6. *Zadeh L.A.* Calculus of Fuzzy Restrictions // Fuzzy Sets and Their Applications to Cognitive and Decision Processes / Ed. by L.A. Zadeh, R.S. Fu and M. Shimura. New York: Academic Press, 1975. Pp. 1–39.
7. *Zadeh L.A.* Fuzzy Sets and Information Granularity // Advances in Fuzzy Sets Theory and Applications. Ed. by M.M. Gupta, R.K. Ragade, R.R. Yager. Amsterdam: North-Holland Publishing Company, 1979. Pp. 3–20.
8. *Zadeh L.A.* Toward a Generalized Theory of Uncertainty (GTU): an Outline // Information Sciences – Informatics and Computer Science. 2005. Vol. 172. № 1–2. Pp. 1–40.
9. *Zadeh L.A.* Fuzzy Logic, Neural Network and Soft Computing // Communications of the ACM. 1994. Vol. 37. № 3. Pp. 77–84.
10. *Аверкин А.Н., Прокопчина С.В.* Мягкие вычисления и измерения // Интеллектуальные системы (МГУ). 1997. Т. 2. Вып. 1–4. С. 93–114.
11. *Tarasov V.B.* General Approaches to the Modelling of Soft Estimates and Beliefs in Strategic Decision Engineering // Proceedings of 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002, Divnomorskoe, Russia, 5–10 September). Los Alamitos CA: IEEE, 2002. Pp. 45–49.

12. *Платон*. Сочинения. В 3-х томах / под ред. В.Ф. Асмуса и А.Ф. Лосева. Т. 1. М., 1968.
13. *Кандрашина Е.Ю., Литвинцева Л.В., Поспелов Д.А.* Представление знаний о пространстве и времени в системах искусственного интеллекта. М.: Наука, 1988.
14. *Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В.* Достоверный и правдоподобный вывод в интеллектуальных системах. М.: Физматлит, 2004.
15. *Нариньяни А.С.* НЕ-факторы и инженерия знания: от наивной формализации к естественной прагматике // Сборник трудов IV-й Национальной конференции по ИИ (КИИ-94, Рыбинск, сентябрь 1994). В 2-х томах. Т. 1. Тверь: АИИ, 1994. С. 9–18.
16. *Konolige K.* A Deduction Model of Belief. San Francisco CA: Morgan Kaufman, 1986.
17. *Fagin R., Halpern J.I.* Belief, Awareness and Limited Reasoning // Artificial Intelligence. 1988. Vol. 34. Pp. 39–76.
18. *Ивин А.А.* Основы теории аргументации. М.: ВЛАДОС, 1997.
19. *Ивин А.А., Никифоров А.Л.* Словарь по логике. М.: ВЛАДОС, 1997.
20. *Тарасов В.Б.* Нечеткость оценок и моделирование субъективных суждений // Психологические механизмы формирования оценочных суждений. Саратов: Изд-во Саратовского госуниверситета, 1989. С. 37–43.
21. *Sugeno M.* Fuzzy Measures and Fuzzy Integrals: a Survey // Fuzzy Automata and Decision Processes / Ed.by M.M. Gupta, G. Saridis, B.R. Gaines. Amsterdam: North-Holland Publishing Company, 1977. Pp. 89–102.
22. Fuzzy Measures and Integrals. Theory and Applications / Ed.by M. Grabish, T. Murofusch, M. Sugeno, J. Kacprzyk. Berlin: Springer-Verlag, 2000.
23. *Dubois D., Prade H.* A Class of Fuzzy Measures Based on Triangular Norms. A General Framework for the Combination of Uncertain Information // International Journal of General Systems. 1982. Vol. 8. № 1. Pp. 43–61.
24. *Shafer G.* A Mathematical Theory of Evidence. Princeton NY: Princeton University Press, 1976.
25. Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д.А. Поспелова. М.: Наука, 1986.

26. *Zadeh L.* Fuzzy Sets as a Basis for a Theory of Possibility // Fuzzy Sets and Systems. 1978. Vol. 1. № 1. Pp. 3–28.

27. *Дюбуа Д., Прад А.* Теория возможностей. Приложения к представлению знаний в информатике / пер. с франц. М.: Радио и связь, 1990.

28. *Luca de D., Termini S.* A Definition of Non-Probabilistic Entropy in the Setting of Fuzzy Sets Theory // Information and Control. 1972. Vol. 20. Pp. 301–312.

29. *Luca de D., Termini S.* Entropy and Energy Measures of a Fuzzy Set // Advances in Fuzzy Sets Theory and Applications / Ed .by M.M. Gupta, R.K. Ragade, R.R. Yager. Amsterdam: North-Holland Publishing Company, 1979. Pp. 321–337.

30. *Higashi M., Klir G.* On Measures of Fuzziness and Fuzzy Complements // International Journal of General Systems. 1982. Vol. 8. Pp. 169–180.

31. *Yager R.R.* On Measures of Specificity // Computational Intelligence: Soft Computing and Fuzzy-Neuro Integration with Applications / Ed. by O. Kaynak, L.A. Zadeh, B. Turksen and I.J. Rudas. Berlin: Springer-Verlag, 1998.

ГЛАВА 5

ТЕОРИЯ ЭНТРОПИЙНЫХ ПОТЕНЦИАЛОВ. СОСТОЯНИЕ, ПЕРСПЕКТИВЫ РАЗВИТИЯ И ПРАКТИЧЕСКОГО ПРИЛОЖЕНИЯ

В.Л. Лазарев

Введение

Теория энтропийных потенциалов (ТЭП), наряду с другими теориями и подходами, является эффективным «инструментом» по описанию и исследованию состояний неопределенности объектов и систем различной природы. Использование ее методов и разработок позволяет осуществлять организацию мониторинга и управления в различных ситуациях и, особенно, в условиях неопределенности. Изложение основных положений и сути ТЭП осуществлено в работах [1–4]. Прикладные особенности, ее практическая направленность и конкретные примеры рассмотрены в фундаментальной коллективной монографии, посвященной организации управления в условиях неопределенности [5]. Там же изложены альтернативные перспективные подходы и направления, проведен их сравнительный анализ и указаны возможные области приложений. ТЭП оказалась востребованной, она развивалась и продолжает развиваться. Результаты последующих разработок были представлены в разрозненных материалах статей в различных журналах и трудах конференций [6–19]. В связи с чем, и возникла необходимость переработки, дополнения и обобщения полученных результатов с целью углубления и совершенствования изложения, а также планирования и организации дальнейших исследований. Данная работа призвана обеспечить решение возникших проблем. Изложение материала осуществлено на основе использования базовых работ [1–5], в соответствии с устоявшимися структурой и традициями. Пополнение и развитие данного направления, в значительной мере, обсуждалось и нашло отражение в трудах ежегодной Международной конференции по мягким вычислениям и измерени-

ям. Эта конференция является признанным «полигоном» для «обкатки» новых разработок, методов и теорий в области создания интеллектуальных технологий мониторинга и управления, а также в смежных областях. Результаты последующих разработок автора, его соратников и учеников нашли отражение в данной работе. В частности, сделаны дополнения по вопросам синтеза и совершенствования информационных моделей, их использования для построения информационных портретов систем, организации мониторинга на основе этих портретов и др.

5.1. Отправные положения

Развитие науки и техники требует углубления исследований свойств объектов и явлений окружающего мира. Поэтому наряду с необходимостью развития технических средств измерений и проведения наблюдений, возрастает актуальность проблемы, связанной с повышением эффективности анализа и обработки исходной информации. В настоящее время проводятся работы по созданию технологий и алгоритмов, позволяющих осуществлять автоматизированную обработку и анализ данных. Современные технологии не являются закрытыми и допускают возможность развития, замены или пополнения новыми блоками методов и алгоритмов анализа и процедур обработки информации.

Наряду с решениями конкретных практических задач, во всем мире проводятся интенсивные исследования, направленные на создание новых подходов, теорий и методов анализа исходных данных. Актуальность проблемы все больше возрастает с повышением требований к «интеллектуализации» и качеству проводимых исследований, прогнозов и организации управления. Потребность в таких исследованиях остро возникает в ситуациях, обусловленных необходимостью учета различных форм и видов неопределенностей в рассматриваемых объектах. В этих случаях возникает необходимость к переходу от «детерминистских» схем и моделей мониторинга и управления к более реальным и объективным, дополнительно учитывающим случайное, априори непредсказуемое проявление различных факторов и субстанций. Тем самым создаются предпосылки «интеллектуализации» процессов обработки наблюдений, прогнозирования и принятия решений по организации управ-

ления в условиях неопределенности. Об актуальности указанных проблем свидетельствует нарастающий поток публикаций.

Понятие «состояние неопределенности» является общим и может быть применено к различным типам систем: детерминированным, вероятностным, нечетким и др. Попытки систематизации и классификации природы возникновения этих состояний предпринимались и продолжают предприниматься отечественными и зарубежными исследователями. Этот процесс далек от окончательного завершения. Тем не менее, представляется возможным выделить основные признаки указанного явления.

Существуют различные варианты трактовок понятия «неопределенность». Тем не менее, можно выделить основные характерные особенности, присущие ему независимо от природы рассматриваемых явлений (техническая, экологическая, биологическая, политическая, социальная и др.) и особенностей протекающих процессов. Состояние системы по рассматриваемым параметрам вследствие различных причин (внесения управляющих воздействий или проявления комплексов каких-либо влияний и связей) характеризуется наличием случайных составляющих. С точки зрения диалектики такое явление естественно и закономерно. Поэтому для его отражения используется специальная философская категория-случайность. Другими словами при наличии любых взаимодействий в системах различной природы возможно появление случайных отклонений от «генеральной» линии развития процесса, что может привести к возникновению нештатных и априори непредсказуемых последствий-проявлений состояний неопределенности.

В детерминированных системах состояние неопределенности также может возникнуть в результате «сужения» рамок рассматриваемой проблемы, когда полностью или частично исключаются связи, влияющие на рассматриваемый параметр. В этом случае для «наблюдателя», находящегося в этих рамках, проявление неучтенных влияний и связей будет представляться элементами состояний неопределенности. Другой причиной возникновения состояний неопределенности может, например, явиться «сжатие» исходных результатов наблюдений с помощью различных аппроксимирующих алгоритмов и моделей. В результате может возникнуть ситуация повышения уровня «неопределенности знаний».

Непредсказуемость, неопределенность результатов реализации того или иного процесса зачастую может быть обусловлена «дефици-

том» исходной информации. Причинами его возникновения могут, например, являться: сложность и дороговизна проведения отдельных измерений, большие затраты времени на проведение наблюдений и др. Такую ситуацию можно трактовать, как своеобразное искажение, «нечеткость» информационного процесса, используемого для получения оценок состояний и выработки решений.

Другой причиной состояния неопределенности по какому-либо параметру является его «рассеяние» в пространственных или временных координатах. Например, разброс значений температур в объеме термической камеры, порождает элемент неопределенности для прогнозов по качеству тепловой обработке изделий и организации управления соответствующими процессами теплообмена. Похожую ситуацию порождает нестабильность температуры во времени на какой-либо из технологических операций в процессе обработки.

В задачах искусственного интеллекта причинами возникновения неопределенностей могут являться:

- несогласованность мнений экспертов;
- неполнота и противоречивость «баз знаний»;
- «нечеткость» или частичная недостоверность исходных данных;
- различные приоритеты «предпочтений» или компонентов целевых функций и пр.

Также существует множество других причин возникновения состояний неопределенности.

5.2. Обзор подходов к описанию состояний неопределенности

Анализ причин возникновения состояний неопределенности позволяет сделать вывод о том, что они являются реалиями в научной и практической деятельности. Их учет, в большинстве случаев, позволяет повысить качество мониторинга, прогнозирования и управления различными объектами и процессами.

Основой этой проблемы является необходимость количественного описания состояний неопределенности в различных ситуациях. Причем для синтеза алгоритмов исследований и логик принятия решений является важным, чтобы мера состояний неопределенности представлялась вещественным числом. Следует отметить, что однозначного решения такой важной проблемы в общем виде не суще-

ствуется. Имеются отдельные альтернативные подходы и направления со своими возможностями, достоинствами и недостатками.

Ниже дается краткий обзор существующих, и перспективных подходов, теорий и методов [20–27].

1. Классические методы спектрального и корреляционного анализа, основанные на рассмотрении функций спектральной плотности – $S_x(\omega)$ и автокорреляционной функции – $R_x(\tau)$ для анализируемой случайной функции.

2. Методы интервальных оценок, основанные на использовании понятий доверительных вероятностей и доверительных интервалов.

3. Методы статистического моделирования и стохастической аппроксимации.

4. Методы, основанные на использовании теории эллиптического оценивания, позволяющие получать гарантированные оценки в целом.

5. Методы теории нечетких или размытых множеств, основанные на использовании так называемых функций принадлежности для описания случайных явлений (*fuzzy logic*).

6. Методы робастного оценивания, основанные на использовании различных подходов: на использовании статистических характеристик, на основе теории чувствительности, на основе теории инвариантности и др.

7. Методы, основанные на использовании экспертных систем с привлечением соответствующих баз знаний (*knowledge based systems*).

8. Методы и технологии искусственных нейронных сетей (*artificial neural networks*).

9. Методы и технологии эволюционного моделирования, основанные на синтезе процессов естественной эволюции системы по ожидаемым сценариям с помощью ЭВМ.

10. Методы когнитивных информационных технологий, в основе которых используются различные процедуры отображения объемных и многомерных информационных массивов в образные когнитивные представления. Эти образы обычно создаются с использованием компьютерной графики, удобны для восприятия и принятия решений в трудно формализуемых ситуациях.

11. Байесовские интеллектуальные технологии (БИТ) и сети. В широком понимании, БИТ это совокупность методов анализа объектов различной природы на основе статистических процедур исследования априорных и апостериорных данных. При этом возможна инте-

грация различных методов определения вероятностных характеристик в условиях априорной неопределенности, «нечеткости» исходных данных и др. Байесовские сети являются одной из моделей «баз знаний» на основе вероятностных характеристик.

12. Методы теории энтропийных потенциалов (ТЭП), основанные на оценивании состояний неопределенности систем по величинам энтропийных потенциалов. Эти оценки являются унифицированными, объективными и достаточно удобно определяемыми. Кроме того, ТЭП и ее методы не являются «изолированными», они могут использоваться в качестве «инструментов» при реализации большинства выше упомянутых подходов и технологий исследований.

Перечисленный перечень направлений, подходов и методов не является исчерпывающим. Важность проблемы породила интерес к проведению подобных исследований. Существуют и другие подходы к описанию состояний неопределенности. Таковыми, например, являются теории Дюбуа-Прада, Демпстера-Шеффера, система INFERNO и др. [28].

Каждый из перечисленных подходов имеет свои преимущества и недостатки. Опыт развития науки и техники, результаты обзора литературных источников позволяют сделать прогноз, что наиболее востребованными будут являться такие подходы к решению задач проведения исследований, прогнозирования и управления в условиях неопределенности, которые позволяют оценивать состояние систем на основе вещественной функции, учитывающей вероятностные свойства рассматриваемых параметров совместно с характеристиками их рассеяния. При этом используемые характеристики должны поддаваться достаточно простому определению, быть компактными.

В значительной мере изложенным требованиям отвечает комплекс понятий энтропийных потенциалов, который введен в рамках ТЭП [1–6]. Эта теория имеет перспективы развития и практического применения.

5.3. Основы теории энтропийных потенциалов

ТЭП использует набор понятий энтропийных потенциалов: энтропийный потенциал (ЭП), комплексный ЭП (КЭП), многомерный КЭП (МКЭП). Эти понятия разработаны на основе принципа «вложения»: понятия более высокого уровня выражаются через величины энтропий-

ных потенциалов предыдущих уровней с введением дополнительных характеристик состояний неопределенности анализируемых параметров. В общем виде всю совокупность понятий энтропийных потенциалов – E можно определить кортежем множеств и отношений вида [1, 5].

$$E = \langle X, N_E, L_{E,Z}, P_E \rangle, \quad (1)$$

где $X = \{\Omega_i\}$, $i=1,2,\dots,m$ – множество параметров, используемых для описания состояния объекта. Количество параметров – m определяет размерность вектора состояния. Это конечное множество может состоять из набора отдельных кластеров $\omega_j(\vec{\xi}) (j=1,2,\dots,k)$, имеющих соответствующие области изменений – D_j . То есть все множество состояний какого-либо i -го параметра $\Omega_i = \bigcup \omega_j(\vec{\xi})$, где $\vec{\xi}$ – вектор факторов, определяющий вариацию величин i -го параметра (например, фактор временных или каких-либо пространственных координат).

N_E – набор отображений для множеств параметров из Ω_i . В качестве таковых в моделях энтропийных потенциалов используются отображения для получения характеристик рассеяния (СКО) – σ_i , базовых значений параметров – X_{ni} , величин энтропийных коэффициентов – K_{ei} , а также, при необходимости, соответствующих весовых коэффициентов – c_i .

$L_{E,Z}$ – набор форм отношений для отображения элементов из N_E в P_E по схеме $N_E \rightarrow P_E$. Z – Номер варианта форм.

P_E – набор понятий, критериев, характеризующих состояния неопределенности элементов из X : ЭП, КЭП, МКЭП.

Приведенное определение (1) допускает дальнейшее развитие и пополнение перечня вводимых понятий для описания состояний неопределенности.

Описание взаимосвязи между понятиями энтропийных потенциалов имеет следующий вид:

$$La_z = \left(\sum_{(i)} \left(c_i |L_{\Delta i}| \right)^z \right)^{\frac{1}{z}} = \left(\sum_{(i)} \left(c_i \frac{\Delta_{ei}}{|X_{ni}|} \right)^z \right)^{\frac{1}{z}} = \left(\sum_{(i)} \left(c_i \frac{\frac{1}{2} e^{H_i}}{|X_{ni}|} \right)^z \right)^{\frac{1}{z}} = \left(\sum_{(i)} \left(c_i \frac{K_{ei} \sigma_i}{|X_{ni}|} \right)^z \right)^{\frac{1}{z}} \quad (2)$$

В определении (2) использованы следующие обозначения:

$L_{\Delta i}$ – КЭП i -ого параметра ($i \in I$).

$$L_{\Delta i} = \frac{\Delta_{ei}}{|X_n|} = \frac{K_{ei} \sigma_i}{|X_n|}, \quad (3)$$

где X_n – величина базового значения, на «фоне» которого рассматривается состояние неопределенности. В качестве базы может быть выбрана величина математического ожидания параметра – m_x , величина его номинального значения и другие величины. Если изменения параметра происходят в окрестности нуля, то в качестве величины X_n могут быть использованы, например, величина диапазона изменения этого параметра, величина предельно-допустимого значения. Необходимость использования модуля величины X_n в выражении (3) обусловлена тем, что в ряде случаев она может принимать отрицательные значения. Такая ситуация, например, может иметь место, когда исследуются состояния неопределенности тепловых режимов объекта в области отрицательных температур. С учетом изложенного и того, что $K_e \geq 0$ и $\sigma \geq 0$, имеем: $L_{\Delta} \geq 0$.

Δ_{ei} – ЭП i -ого параметра. Согласно определению [1], энтропийным потенциалом параметра x называется половина диапазона изменения ограниченного распределения, имеющего такую же энтропию H_x , что и закон распределения этого параметра. В качестве базы для нахождения величины ЭП должно быть выбрано распределение с ограниченным диапазоном изменения $[-\Delta_e, \Delta_e]$. То есть:

$$x \in [-\Delta_e, \Delta_e]. \quad (4)$$

В этом случае плотность распределения вероятностей будет зависеть от величины Δ_e , то есть:

$$p(x) = p(x, \Delta_e). \quad (5)$$

(В качестве $p(x)$ целесообразно использовать функцию симметричную относительно центра диапазона $[-\Delta_e, \Delta_e]$.) Очевидно, что величина энтропии базового распределения также будет зависеть от величины Δ_e в соответствии с (5). Приравнивая энтропию рассматриваемого параметра с произвольным законом распределения $p(x)$

$H_x = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx$, энтропии базового распределения с ограниченным диапазоном изменения параметра $H_x(\Delta_e)$ получим:

$$H_x = H_x(\Delta_e). \quad (6)$$

Разрешая уравнение (6) относительно Δ_e получим выражение для нахождения величины ЭП в общем виде:

$$\Delta_e = F(H_x). \quad (7)$$

Очевидно, что величина Δ_e будет зависеть от вида выбранного базового закона распределения. В такового целесообразно выбрать закон равномерной плотности вероятностей с $p(x) = \frac{1}{2\Delta_e}$. В этом случае будем иметь:

$$H_x = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx = - \int_{-\infty}^{\infty} \frac{1}{2\Delta_e} \ln \frac{1}{2\Delta_e} dx = - \int_{-\Delta_e}^{\Delta_e} \frac{1}{2\Delta_e} \ln \frac{1}{2\Delta_e} dx = \ln 2\Delta_e.$$

Откуда получаем выражение для Δ_e :

$$\Delta_e = \frac{1}{2} e^{H_x}. \quad (8)$$

Аналогичным образом можно получить выражения для величин ЭП на базе других законов распределений и установить между ними взаимосвязь, которая осуществляется с помощью так называемых коэффициентов перехода. Эти вопросы подробно рассмотрены в работах [1, 2].

Частным понятием величины ЭП в виде (8) является понятие энтропийного значения погрешности средств измерений [1, 29]. Оно используется в измерительной технике для описания метрологических свойств различных приборов и измерительных систем.

Величина ЭП имеет размерность рассматриваемого параметра. Поэтому, для каждого конкретного закона распределения, ее можно выразить как масштабное изображение величины среднеквадратического отклонения (СКО) – σ , имеющей такую же размерность. Соот-

ветствующий коэффициент носит название энтропийного коэффициента и обозначается K_e . Следовательно:

$$\Delta_e = K_e \sigma. \quad (9)$$

Выражение (9) позволяет выразить состояние неопределенности параметра через характеристику его рассеяния – σ и коэффициент K_e , характеризующий «дестабилизирующие» свойства его закона распределения. Для ряда типовых законов распределений выражения для нахождения K_e могут быть получены аналитическим путем с использованием выражений (7) и (9). В ряде случаев – на основе экспериментальных данных с использованием выражений (8) и (9) в виде:

$$K_e = \frac{\Delta_e}{\sigma} = \frac{e^{H_x}}{2\sigma}. \quad (10)$$

Возможны иные способы определения K_e : на основе аналогий, с использованием специальных методик и др. [1, 29]. Следует отметить, что помимо выражения (9) величину ЭП можно также выразить и через какие-либо другие характеристики распределения. Например, ее можно представить как масштабное изображение половины величины размаха распределения a в виде $\Delta_e = K_e(a) \cdot a$, где $K_e(a)$ энтропийный коэффициент «привязанный» к размаху выборки. Возможны и другие представления, однако, для конкретики и однозначности проведения исследований будем пользоваться выражением (9).

c_i – весовые коэффициенты, характеризующие значимость каждого i -ого параметра при описании состояния системы, $c_i \geq 0$, $(\sum_{(i)} c_i = 1)$.

Значения весовых коэффициентов могут быть определены аналитически, на основании методов экспертных оценок, на основании приоритетов, определяемых существующими стандартами и технологическими инструкциями. Так, например, при оценивании состояния неопределенности какого-либо параметра на основании ряда неравноточных независимых наблюдений, весовые коэффициенты для результатов измерений в отдельных наблюдениях будут обратнопропорциональны значениям соответствующих дисперсий.

Z – номер варианта критерия ($z=1 \dots n$), где n – количество вариантов. При $z=1$ получаем вариант критерия La_1 :

$$La_1 = \sum_{i=1}^m (c_i \cdot L_{\Delta i}). \quad (11)$$

При $z=2$ – вариант La_2 :

$$La_2 = \sqrt{\sum_{i=1}^m (c_i \cdot L_{\Delta i})^2}. \quad (12)$$

В геометрическом представлении величина La_2 является модулем или длиной m -мерного вектора, составляющими которого являются величины КЭП отдельных параметров системы в масштабе их весовых коэффициентов. Величина La_1 является суммой длин модулей этих составляющих. Поэтому имеет место условие: $La_1 \geq La_2$. Причем равенство имеет место в случае, когда $m=1$. Выбор варианта критерия является прерогативой пользователя. Критерий La_2 дает оценку состояния неопределенности менее зависимую от значения величины размерности системы m , чем критерий La_1 . Поэтому, при достаточно больших значениях величины m ($m \geq 5$) предпочтительней является критерий La_2 . Возрастание величин La_z , в обоих вариантах, свидетельствует о повышении уровня состояния неопределенности системы и, наоборот. Согласно определениям, величины La_z и L_{Δ} являются безразмерными, что позволяет использовать их в качестве критериев подобия при исследовании состояний неопределенности различных систем.

Очевидно, что при $m=1$ величина La_z вырождается в величину L_{Δ} . Если базовое значение параметра X_n в конкретной ситуации является постоянным, то величина L_{Δ} будет являться безразмерным масштабным изображением величины энтропийного потенциала параметра – Δ_e . Если еще пренебречь и трансформацией вариативных свойств закона распределения в процессах эволюции или управления системой, что соответствует условию $K_e = const$, то величина L_{Δ} будет являться масштабным изображением величины СКО. В этом частном случае исследование состояний системы может быть осуществлено с использованием известных методов дисперсионного анализа. В другом частном случае, когда $K_e = const$, а в качестве базового значения выбрана величина математического ожидания параметра, т.е. $X_n = m_x$, то величина L_{Δ} будет являться масштабным изображением коэффициента вариации

параметра V_x . Этот коэффициент также используется в статистике для характеристики разброса, неопределенности значений параметра.

Таким образом, понятия величин ЭП являются взаимосвязанными и основаны на принципах «вложения»: понятия более высокого уровня выражаются через величины энтропийных потенциалов предыдущих уровней. Процедура перехода между вариантами моделей энтропийных потенциалов осуществляется путем исключения или дополнительного учета характеристик неопределенности (варианта многопараметрической модели, базового значения, энтропийного коэффициента и др.). Схема преобразования моделей состояний неопределенности, основанных на использовании комплекса понятий ЭП, изображена на рис. 1.

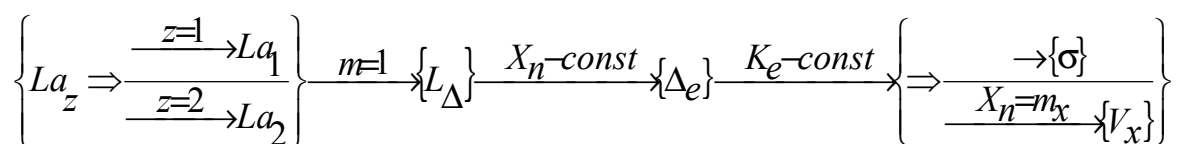


Рис. 1. Схема преобразования моделей состояний неопределенности на основе комплекса понятий энтропийных потенциалов

Изображенная на рис. 1 схема иллюстрирует процедуру перехода от более полных вариантов моделей состояний неопределенности к более простым. Также возможен переход и в обратном направлении за счет введения или учета вышеуказанных характеристик состояний неопределенностей в модели предыдущих уровней.

Предложенный набор понятий величин энтропийных потенциалов позволяет описать разные «уровни» состояний неопределенности в зависимости от специфики поставленной задачи. Получаемые модели являются не противоречивыми, компактными и удобными для практического применения. С их помощью удастся описать состояния неопределенности различных объектов и систем единым комплексом, основанным на использовании наглядных и информативных характеристик: K_e , X_n , σ . Эти характеристики поддаются достаточно простому определению. В случаях, когда они определяются на основании результатов наблюдений, для их определения требуется объем данных на порядок меньший, чем для определения энтропии [1–5, 29, 30].

Контроль состояния неопределенности системы, в общем случае, может быть осуществлен на основании величины многомерного комплексного энтропийного потенциала системы La_z . Для этого сначала

необходимо выбрать вариант критерия La_z , задавшись числом z , определить весовые коэффициенты для всех параметров системы, рассчитать величину La_z , и по значению этой величины оценить состояние системы. Управление состоянием неопределенности системы осуществляется по величине La_z путем целенаправленного изменения характеристик отдельных параметров системы: K_{ei} , X_{ni} , σ_i ($i \in I$). Так, например, изменение величины K_e при управлении может быть осуществлено путем трансформации закона распределения параметра. Реализация такой задачи может быть осуществлена за счет изменения параметров системы, введения специальных корректирующих звеньев и др. Практическая сторона этих вопросов подробно рассмотрена в работах [1, 4, 31, 32].

Проиллюстрируем применение изложенного подхода на примере системы контроля и управления температурно-влажностным режимом в камере, предназначенной для климатических испытаний различной аппаратуры, механизмов, материалов, покрытий и др. [5]. Состояние такой системы при проведении испытаний характеризуется двумя параметрами ($m=2$): температурой- $y_1(^{\circ}C)$ и относительной влажностью воздуха- $y_2(\%)$ в заданных точках контроля. Для данного числа параметров ($m < 5$) выбираем вариант критерия La_1 (11). Оба параметра одинаково значимы для проведения испытаний, поэтому весовые коэффициенты априори приняты равными: $c_1=c_2=1/2$. Поддержание температурно-влажностного режима в камере осуществляется с помощью соответствующих контуров регулирования, качество работы которых характеризуется значениями величин текущих отклонений параметров от базовых значений с соответствующими вероятностями (то есть соответствующими законами распределений). Другими словами, качество работы контуров регулирования оказывает влияние на состояние неопределенности температурно-влажностного режима камеры. Базовые значения параметров в каждом климатическом режиме являются уставками соответствующих регуляторов. На основании обработки результатов экспериментальных данных по регулированию параметров в одном из рабочих режимов были получены следующие характеристики. По температуре: $X_{n1}=40^{\circ}C$; $K_{e1}=2.01$; $\sigma_1=0.8^{\circ}C$. По относительной влажности: $X_{n2}=85\%$; $K_{e2}=1.8$; $\sigma_2=1.7\%$. Для этого режима величины комплексных энтропийных потенциалов будут соответственно равны (3): $L_{\Delta 1} = (2.01 \cdot 0.8)/40 \approx 0.04$; $L_{\Delta 2} = (1.8 \cdot 1.7)/85 \approx 0.036$. Состояние неопределенности данного температурно-влажностного режима в рассматриваемой систе-

ме определяем с помощью величины многомерного комплексного энтропийного потенциала $La_1=(1/2) \cdot L_{\Delta 1}+(1/2) \cdot L_{\Delta 2}=0,038$. С целью повышения качества поддержания температурно-влажностного режима (т.е. уменьшения его состояния неопределенности) была проведена коррекция обоих контуров регулирования путем изменения настроечных параметров регуляторов. Здесь необходимо отметить следующие обстоятельства. Например, для системы регулирования, основанной на использовании промышленного ПИД (пропорционально-интегро-дифференцирующего) регулятора с передаточной функцией $W(p)=K(1+T_1p+1/(T_2p))$, настроечными параметрами регулятора (т.е. параметрами, которые можно изменять в процессе эксплуатации и для изменения которых имеются специальные аппаратные средства) являются K – коэффициент передачи регулятора, T_1 – время предварения, T_2 – постоянная интегрирования. Используя эти настройки можно решать различные задачи по управлению процессом регулирования: изменять запас устойчивости системы, изменять вид и показатели качества процесса регулирования (например, получать аperiodический или колебательный переходной процесс, изменять величины перерегулирования, ошибок и др.). В частности, если управление процессом регулирования осуществлять по величине комплексного энтропийного потенциала – L_{Δ} , то, изменяя величины K , T_1 и T_2 будем изменять передаточную функцию регулятора а, следовательно, и всего контура регулирования. В результате будут изменяться динамические свойства системы, ее частотная характеристика и степень влияния различных возмущений на регулируемый параметр. Это приведет к изменению спектрального состава динамической составляющей регулируемого параметра и закона распределения величин его отклонений от уставного или базового значения. В конечном счете, изменятся, в различных соотношениях, величины σ и K_e , а, следовательно, и величины энтропийного потенциала и комплексного энтропийного потенциала. В скорректированном варианте системы регулирования характеристики параметров получились следующими. По температуре: $X_{n1}=40^{\circ}\text{C}$; $K_{e1}=1.95$; $\sigma_1=0.5^{\circ}\text{C}$. По относительной влажности: $X_{n2}=85\%$; $K_{e2}=1.92$; $\sigma_2=1.2\%$. Величины комплексных энтропийных потенциалов будут соответственно равны: $L_{\Delta 1}=(1.95 \cdot 0.5)/40 \approx 0.024$; $L_{\Delta 2}=(1.92 \cdot 1.2)/85 \approx 0.027$. Состояние неопределенности температурно-влажностного режима в этой ситуации характеризуется новым значением величины $La_1 \approx 0.026$. Из приведенных данных видно, что в результате реализации таких действий, состояние неопределенности системы по рас-

смаатриваемым параметрам y_1 и y_2 уменьшилось, или, другими словами, возросло качество поддержания температурно-влажностного режима. Рассмотренный цикл изменения настроечных параметров регулятора можно трактовать как этап итеративной процедуры синтеза системы управления по методу минимизации энтропийного потенциала. При этом существует множество алгоритмов и планов поиска решений [21, 27].

В приведенном примере состояние системы по выбранным параметрам рассматривается в заданных точках с учетом их изменения во времени в процессе регулирования. Аналогично можно контролировать состояние системы в пространстве, то есть в объеме камеры. В этом случае необходимо рассматривать распределения значений параметров y_1 и y_2 в разных точках объема камеры, обусловленные неоднородностью температурного и влажностного полей. Таким же образом можно контролировать изменение состояния неопределенности системы, вызванное изменением базовых значений параметров, соответствующим переходу камеры на другой климатический режим и др.

Аналогичный подход был применен в работе [3] для исследования эволюций температурного режима г. Санкт-Петербурга в период с 1865 по 1993 гг. На основании полученных результатов были сделаны выводы и прогнозы об изменении климата в регионе. Такой же подход оказался продуктивным при организации исследований социальных процессов, экономических, биологических и др. систем [33-38].

Ниже рассматриваются основные перспективные направления проведения исследований с использованием понятий энтропийных потенциалов и определяющих их характеристик.

Достоинствами ТЭП является то обстоятельство, что с использованием ее методов представляется возможным исследовать состояния различных систем с помощью фазовых портретов в пространстве характеристик ЭП, а также на основе информационных портретов или по их информационному «следу». Изложению этих вопросов посвящены следующие подразделы.

5.4. Исследование систем на основе портретов в пространстве характеристик энтропийных потенциалов

Суть проведения таких исследований основана на использовании понятия фазового пространства и состоит в следующем.

Понятие фазового пространства используется для качественного анализа процессов эволюций систем различной природы. В обобщенном смысле, понятие «фаза» трактуется как этап, стадия изменения состояния системы. Под фазовым пространством подразумевается, декартова система координат, в качестве которых рассматриваются величины, характеризующие какие-либо свойства системы [27, 39]. При исследовании динамики в качестве таковых, чаще всего, используется выходная величина системы и $(n-1)$ ее производные. При этом, в геометрической интерпретации, состояние системы в любой момент времени будет описываться положением т. н. изображающей точки в фазовом пространстве. Изменение состояний системы будет характеризоваться перемещениями этой точки, которые образуют соответствующую характерную траекторию. Совокупность фазовых траекторий, соответствующих различным режимам функционирования системы или эволюционным процессам, образует ее фазовый портрет. Фазовые портреты могут рассматриваться как элементы когнитивной графики. Они позволяют получать представление о процессах эволюции систем, видах переходных процессов и др. [4, 11, 39].

В представленном подходе, в качестве координат фазового пространства предлагается использовать величины, входящие в выражения для энтропийных потенциалов анализируемых параметров системы [1–5, 10–12]. Наряду с этим также возможно использование величин энтропийных потенциалов и их производных, что позволяет отображать и исследовать динамику изменения состояний неопределенности. Использование величин σ , K_e и X_n в качестве координат фазового пространства позволяет получать компактные и информативные портреты эволюции состояний неопределенности, характеризующиеся высоким уровнем «сжатия» исходной информации. Такой подход имеет перспективы дальнейшего развития.

Методологические аспекты построения и особенности портретов в пространстве характеристик энтропийных потенциалов, в общем виде, состоят в следующем.

Согласно обобщенному определению ЭП (2) состояние неопределенности любой системы по m параметрам определяется набором $3m$ характеристик (K_{ei} , X_{ni} , σ_i ; $i=1,2,\dots m$). В геометрической интерпретации это состояние можно изобразить точкой в $3m$ -мерном пространстве указанных характеристик. Изменение состояния неопреде-

ленности будет характеризоваться перемещением изображающей точки в этом пространстве. Неизменности состояния неопределенности системы будет соответствовать перемещение изображающей точки по $3m$ -мерной поверхности, определяемой условием:

$$La_z = C = const. \quad (13)$$

Такая поверхность является изопотенциальной поверхностью в пространстве параметров МКЭП. Изменению состояния неопределенности системы будет соответствовать переход изображающей точки на другую изопотенциальную поверхность, определяемую другим соответствующим значением величины C . Используя определение (13) можно сформулировать условие неизменности состояния неопределенности системы в общем виде:

$$d(La_z) = 0. \quad (14)$$

В геометрической интерпретации условие (14) будет соответствовать перемещению изображающей точки по изопотенциальной поверхности. При использовании, например, критерия La_1 условие (14) примет вид:

$$\begin{aligned} d(La_1) &= d\left(\sum_{(i)} c_i L_{\Delta i}\right) = \sum_{(i)} c_i d(L_{\Delta i}) = \sum_{(i)} c_i d\left(\frac{K_{ei} \sigma_i}{X_{ni}}\right) = \\ &= \sum_{(i)} c_i \left(\frac{K_{ei} d\sigma_i}{X_{ni}} + \frac{\sigma_i dK_{ei}}{X_{ni}} - \frac{K_{ei} \sigma_i dX_{ni}}{X_{ni}^2}\right) = 0 \end{aligned} \quad (15)$$

В частном случае, когда $m=1$, т.е. когда состояние неопределенности системы определяется одним параметром условие (14) упрощается до выражения:

$$\begin{aligned} \frac{K_e d\sigma}{X_n} + \frac{\sigma dK_e}{X_n} - \frac{K_e \sigma dX_n}{X_n^2} &= \frac{\left(K_e d\sigma + \sigma dK_e - \frac{K_e \sigma dX_n}{X_n}\right)}{X_n} = \\ &= \frac{K_e d\sigma + \sigma dK_e - K_e \sigma d(\ln X_n)}{X_n} = 0. \end{aligned} \quad (16)$$

Выражение (16), в геометрической интерпретации, описывает перемещение изображающей точки по изотропной поверхности в 3-х мерном пространстве координат: K_e , X_n , σ . С учетом того, что $X_n \neq 0$, выражение (16) может быть преобразовано к виду:

$$K_e d\sigma + \sigma dK_e - K_e \sigma d(\ln X_n) = 0. \quad (17)$$

В выражении (15) величина дифференциала dK_{ei} характеризует изменение энтропийных или «дестабилизационных» свойств закона распределения i -ого параметра; величина дифференциала $d\sigma_i$ – изменение величины СКО i -ого параметра; величина дифференциала $d(\ln X_{ni})$ – изменение величины базового значения i -ого параметра в логарифмическом масштабе ($i \in I$). Эти дифференциалы, характеризующие изменения соответствующих параметров энтропийных потенциалов, могут быть использованы для образной иллюстрации отдельных составляющих состояний неопределенности. Величина σ характеризует усредненную мощность всего спектра гармоник динамической составляющей рассматриваемого параметра, ее можно рассматривать, как «энергетическую» компоненту этих состояний. Величина X_n согласно определению характеризует «фоновую» или базовую составляющую. Величина K_e согласно выражению (10) связана с энтропией системы а, следовательно, ее изменение связано с количеством информации I_L , порожденной изменением состояния неопределенности. Поэтому естественным будет считать, что величина dK_e характеризует проявление информационных или «интеллектуальных» составляющих этих состояний. Используя предложенную трактовку, представляется возможным наглядно оценить и смоделировать влияние энергетической, базовой и интеллектуальной составляющих на изменение состояния неопределенности какой-либо системы.

В еще более частном случае когда, например, величина базового значения является неизменной, т.е. $X_n = const$, выражение (17) упрощается до вида:

$$K_e d\sigma + \sigma dK_e = 0. \quad (18)$$

В этом случае изопотенциальная поверхность «выродится» в кривую $\Delta_e = K_e \sigma = C = const$ в системе координат σ , K_e на плоскости. Эта кривая называется линией постоянного энтропийного потенциала или, более кратко, изотропой. Изменение состояний неопределенно-

сти системы будет характеризоваться перемещением изображающей точки на указанной плоскости, которая будет описывать некоторую кривую. Совокупность таких кривых, соответствующих различным вариантам начальных условий и эволюции систем называется портретом ЭП или, более кратко, энтропийным портретом. Свойства изотроп и энтропийных портретов подробно исследованы в работах [1–4].

Если пойти на дальнейшее упрощение описания состояния неопределенности и пренебречь изменением вариативных свойств закона распределения параметра, т.е. считать, что $K_e = \text{const}$ в выражении (18), то условие постоянства состояния неопределенности сведется к неизменности величины СКО – σ . То есть условие (18) «упростится» до условия:

$$d\sigma = 0. \quad (19)$$

В этом случае исследование состояний неопределенности «сведется» до уровня методов дисперсионного анализа.

Очевидно, что набор условий (14)–(19) является не противоречивым и также соответствует вышеуказанным принципам «вложения».

Портреты систем в пространстве характеристик энтропийных потенциалов являются весьма «информативными» и позволяют просто и наглядно проиллюстрировать изменение состояния неопределенности любой системы. Ниже приведены наглядные примеры и соответствующие иллюстрации.

5.5. Информационные портреты

Используя понятия ЭП, представляется возможным и удобным описать количественную взаимосвязь между количеством информации, порожденной изменением состояния неопределенности системы, и соответствующими величинами ЭП и их характеристиками.

В качестве общего примера рассмотрим ситуацию, когда состояние неопределенности системы оценивается по величине КЭП. В этом случае, изменение состояния неопределенности системы на каком-либо этапе ее существования будет характеризоваться соотношением величин $L_{\Delta 1}$ и $L_{\Delta 2}$, соответствующим ее состояниям в начале и конце этого этапа. (Изменение состояния неопределенности, например, может быть обусловлено различными эволюционными процессами или внесением каких-либо воздействий и др.). Количественная характеристика этого из-

менения может быть представлена в виде отношений соответствующих величин КЭПов и выражена через их составляющие в следующем виде:

$$\frac{L_{\Delta 1}}{L_{\Delta 2}} = \frac{\Delta_1}{X_{n1}} \cdot \frac{X_{n2}}{\Delta_2} = \frac{X_{n2}}{X_{n1}} \cdot \frac{\Delta_1}{\Delta_2} = \frac{X_{n2}}{X_{n1}} \cdot \frac{e^{H_1}}{e^{H_2}} = \frac{X_{n2}}{X_{n1}} \cdot e^{H_1-H_2} = \frac{X_{n2}}{X_{n1}} \cdot e^I. \quad (20)$$

При преобразованиях в выражении (20) использовались следующие обозначения и представления.

Величины энтропийных потенциалов Δ_1 и Δ_2 преобразованы в соответствии с выражением (8), H_1 и H_2 – энтропии системы по рассматриваемому параметру в начале и конце анализируемого этапа, соответственно. $I = H_1 - H_2$ – количество информации, порожденное изменением состояния неопределенности системы.

Используя выражение (20) несложно получить искомое выражение для величины I .

$$I = \ln \left(\frac{L_{\Delta 1}}{L_{\Delta 2}} \cdot \frac{X_{n1}}{X_{n2}} \right) = \ln \frac{\Delta_1}{\Delta_2}. \quad (21)$$

Полученный результат (21) позволяет описать «информационный след» системы используя понятия энтропийных потенциалов. Очевидно, что полученная модель является компактной и удобной для практических расчетов, в частности, хотя бы потому, что для вычисления I требуется меньший объем выборки наблюдений, чем в «классическом» случае, основанном на определении убыли энтропии, как, например, в выражении (20). Исследование свойств информационного поля систем на основании величин энтропийных потенциалов произведено строго и полно в работах [1, 2] с помощью набора соответствующих выкладок и теорем. Не вдаваясь в теоретические «тонкости» необходимо отметить лишь ряд особенностей данного подхода, представляющих интерес для практических исследований.

1. Количество информации I , порожденное изменением состояния неопределенности системы «инвариантно» относительно величин базовых значений ее параметров. Поэтому для проведения соответствующих расчетов достаточно ограничиться использованием величин ЭП. Это утверждение вытекает непосредственно из окончатель-

ного варианта выражения (21). В работе [1] доказательство этого результата представлено в виде Теоремы 3.

2. Величина I может принимать как отрицательные, так и положительные значения ($-\infty < I < \infty$). Положительные значения величины I соответствуют уменьшению, понижению уровня состояния неопределенности системы в конце рассматриваемого этапа и наоборот. Ситуация когда $I=0$ соответствует неизменности уровня неопределенности. Сказанное непосредственно вытекает из выражения (21). Тот факт, что величина I может принимать отрицательные значения, не противоречит положениям «классической» теории информации, основанной на использовании понятий априорной и апостериорной энтропий. При этом полагается, что апостериорная энтропия всегда меньше априорной, так как получается после устранения какого-либо элемента неопределенности системы. Таковым, например, может быть проведение цикла измерений, прием «уточняющего» сигнала и пр. В данном, более общем случае, допустимо увеличение «последующей» энтропии, т.е. в конце рассматриваемого этапа, что подразумевает возможность проявления дестабилизирующих событий и факторов. Поэтому введенное на основании выражения (21) понятие информации является обобщающим, включающим в себя «классическое» определение, и для исключения неоднозначности трактовки, в дальнейшем будет обозначаться I_L .

3. Одним из достоинств обобщенного понятия информации I_L является то обстоятельство, что с его помощью представляется возможным описать взаимосвязь между информационными, энергетическими и материальными субстанциями. Действительно, из выражений (9) и (21) следует:

$$I_L = \ln \frac{\Delta_1}{\Delta_2} = \ln \frac{K_{e1}\sigma_1}{K_{e2}\sigma_2} = \ln k_{ke} + \ln \frac{\sigma_1}{\sigma_2} = \ln k_{ke} + \ln \frac{\sqrt{D_1}}{\sqrt{D_2}}, \quad (22)$$

где $k_{ke} = \frac{K_{e1}}{K_{e2}}$ – коэффициент преобразования закона распределения параметра. Понятие k_{ke} впервые было введено в работе [4]. Значения k_{ke} в ряде случаев могут быть определены теоретически, исходя из физического смысла и др., а для некоторых типовых ситуаций могут быть заранее вычислены и табулированы. Рекомендации и методики определения значений величины k_{ke} изложены, например, в работе [1]. Ве-

личина k_{ke} характеризует трансформацию вариативных свойств закона распределения параметра, что является одним из аргументов в задачах «интеллектуального» мониторинга. Поэтому слагаемое $\ln k_{ke}$ целесообразно трактовать, как интеллектуальную составляющую величины I_L . D_1, D_2 – дисперсии параметра системы в начале и конце рассматриваемого этапа. Величина дисперсии характеризует усредненную мощность всего спектра гармоник динамической составляющей рассматриваемого параметра. ($D = \sigma^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) d\omega$, где $S(\omega)$ – функция спек-

тральной плотности центрированной составляющей, которая описывает распределение среднего значения мощности случайного процесса по гармоникам.) Ее можно рассматривать как энергетическую составляющую явления или процесса. Формами проявления связей между энергетической и «материальной» составляющими являются фотоэлектрические, квантовые и др. взаимодействия, а также известное явление «дефекта массы», которое учитывается и используется в различных отраслях науки и техники [40–43]. Поэтому модель (22) потенциально может являться одной из составляющих при описании взаимосвязей в комплексе субстанций: информация-энергия – вещество.

4. На основе обобщенного понятия информации I_L представляется возможным построить «информационный портрет» или информационный «след» системы, характеризующий изменение ее состояния неопределенности в зависимости от какого-либо фактора, например, времени, пространственной метрики. Так если выбран фактор времени τ , то при построении зависимости $I_L(\tau)$ необходимо задаться конечными интервалами времени $\Delta\tau_i$ ($i \in I$) для соответствующих этапов. После чего для каждого из этапов, в соответствии с выражением (21), вычисляется значение величины I_{Li} . Для наглядности, информационный портрет системы может быть представлен в системе координат τ, I_L .

Для иллюстрации изложенного подхода на рис. 2 представлен вариант качественного вида информационного портрета, характеризующего состояния неопределенности температурного режима холодильной камеры. В силу ряда конструктивных особенностей конкретного образца камеры (организации циркуляции потоков холодного воздуха, схемы расположения панелей охлаждения и др.) имеет место разброс температур в объеме камеры, что и обуславливает состояние неопределенности температурного режима для пользователя в процессе ее эксплуатации.

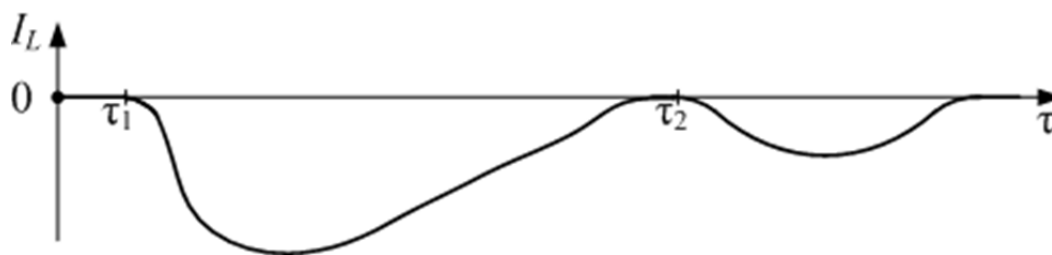


Рис. 2. Информационный портрет состояния неопределенности температурного режима холодильной камеры

До момента времени τ_1 в камере существовал стационарный режим состояния неопределенности температурного поля, т.е. $\Delta_{e0} = \text{const}$. В этой ситуации, согласно (21), $I_L = 0$. В момент времени τ_1 имело место возмущение температурного режима, обусловленное открыванием дверей камеры для загрузки значительной партии продукции. В результате притока тепла, внесенного более теплым наружным воздухом и тепла, внесенного массой этой партии, произошло возрастание состояния неопределенности температурного поля, т.е. увеличение значения величины ЭП ($\Delta_{e1} > \Delta_{e0}$). Вследствие этого обстоятельства, после окончания этапа загрузки, величина I_L приняла отрицательное значение. В дальнейшем, в результате работы системы охлаждения, произошла стабилизация состояния неопределенности температурного поля ($I_L = 0$). Как показано на рис. 2, переход в это состояние был осуществлен по аperiодической траектории. Похожая ситуация имеет место когда возмущение поступает по каналу изменения уставки регулятора температуры (момент времени τ_2 на рис. 2), т.е. при переводе камеры на другой режим охлаждения. В этом случае также имеет место изменение состояния неопределенности температурного поля, но в приведенном примере в меньшей степени, что и отражено «меньшим» всплеском соответствующего фрагмента кривой $I_L(\tau)$ на рис. 2. Следует отметить, что «размеры» обоих всплесков могут варьироваться и, в общем случае, зависят также от значений величин указанных возмущений. Целесообразность представления качественного вида кривой $I_L(\tau)$ на рис. 2 обусловлена намерением обобщения результатов соответствующих исследований.

5. Так как введенное выше понятие позволяет получать количественные оценки изменений состояния неопределенности систем, то с его использованием появляется возможность определять устойчивость этих состояний, используя классические подходы и определения. Так,

например, состояние неопределенности температурного режима вышерассмотренной камеры является устойчивым, так как после устранения возмущения, вызвавшего изменение этого состояния (завершение операции загрузки-выгрузки продукции), система возвращается к прежнему уровню неопределенности. Или, другими словами, переходной процесс в системе является «затухающим», что наглядно проиллюстрировано на рис. 2 с использованием величины I_L .

6. Представляется возможным осуществить дальнейшее развитие выражения (22) с целью получения репрезентативных информационных моделей для различных ситуаций [6–8, 12]. Представим выражение (22) в следующем виде:

$$I_L = \ln \frac{\Delta_1}{\Delta_2} = \ln \frac{K_{e1}\sigma_1}{K_{e2}\sigma_2} = \ln k_{ke} + \ln \frac{\sigma_1}{\sigma_2} = I_I + I_P. \quad (23)$$

В выражении (23) введены обозначения: величина $I_I = \ln k_{ke}$ – интеллектуальная составляющая информации, величина $I_P = \ln \frac{\sigma_1}{\sigma_2}$ – энергетическая составляющая информации. Пояснение смысла этих обозначений приведено в комментариях к выражению (22).

Информационная модель (23), в ряде ситуаций, допускает дальнейшие упрощения [7, 8]. Если известно, что, в силу каких-либо обстоятельств, величина I_I много меньше величины I_P ($I_I \ll I_P$), то величиной I_I в выражении (23) можно пренебречь и «информационный след» системы на этом этапе может быть определен по величине энергетической составляющей информации. То есть:

$$I \approx I_P = \ln \frac{\sigma_1}{\sigma_2} = \ln \sigma_1 - \ln \sigma_2. \quad (24)$$

В частности, величина I_I может быть близка к нулю в случае, когда $K_{e1} \approx K_{e2}$. Это означает, что дестабилизационные свойства закона распределения параметра на рассматриваемом эволюционном этапе не претерпели существенных изменений. Также выражение (24) будет справедливо, если пренебречь изменением его закона распределения.

Если $I_I \gg I_P$, то величина I может быть приблизительно определена по величине I_I :

$$I \approx I_I = \ln \frac{K_{e1}}{K_{e2}} = \ln k_{ke} = \ln K_{e1} - \ln K_{e2}. \quad (25)$$

Такая ситуация возможна когда $\sigma_1 \approx \sigma_2$. В этом случае $I_P \approx 0$. Это соответствует тому, что на данном этапе эволюции системы, величина СКО параметра существенно не изменилась.

Для выбора варианта информационной модели I необходимо задать параметр, количественно описывающий соотношение величин I_I и I_P . В качестве такового предложен коэффициент соотношения информации – η [8]:

$$\eta = \frac{|I_P|}{|I_I|}. \quad (26)$$

Знаки абсолютных величин в выражении (26) необходимы для обеспечения постоянства знака величины η , поскольку величины I_I и I_P , согласно определениям (24) и (25), могут принимать как положительные, так и отрицательные значения.

Выбор конкретного варианта информационной модели необходимо осуществлять исходя из требований к ее адекватности, путем задания граничного уровня значения величины $\eta = C$. Исходя из опыта использования информационных моделей на практике, в качестве такового можно рекомендовать значение C , находящееся в пределах от двух до пяти. Отсюда, если $\eta \leq 1/C$, то определение величины I следует осуществлять из выражения (25). Если $\eta \geq C$, то та же задача должна решаться с использованием выражения (24). В случае если $1/C < \eta < C$, то следует использовать выражение (23).

Выбор значения величины C может быть осуществлен из следующих соображений. Определим значимость или долю каждой из составляющих информации в общей сумме абсолютных величин этих составляющих. Обозначим ее θ и выразим через величину η , используя выражение (26), в следующем виде. Для величины I_I :

$$\theta_I = \frac{|I_I|}{|I_I| + |I_P|} = \frac{|I_I|}{|I_I| + \eta |I_I|} = \frac{1}{1 + \eta}, \text{ или } \eta = \frac{1 - \theta_I}{\theta_I}. \quad (27)$$

Для величины I_P :

$$\theta_P = \frac{|I_P|}{|I_I| + |I_P|} = \frac{|I_P|}{\frac{1}{\eta}|I_P| + |I_P|} = \frac{\eta}{1 + \eta}, \text{ или } \eta = \frac{\theta_P}{1 - \theta_P}. \quad (28)$$

Теперь задавшись уровнем значимости для вариантов информационных моделей в виде конкретного значения величины θ , с использованием выражений (27) и (28), можно определить граничные уровни величины η в виде значений C и $1/C$. Так, например, задавшись 80% уровнем значимости в виде условия $\theta = 0.8$, получим $C=4$ и, соответственно, $1/C=0.25$. Изложенный подход допускает наглядную графическую иллюстрацию, приведенную на рис. 3. Здесь представлены графики зависимостей величин θ_I и θ_P от величины η . Из определений следует, что эти кривые пересекаются в точке с координатами (1, 0.5). Уровень значимости полной информационной модели (23) $\theta = \theta_I + \theta_P = 1$ для любых значений величины η . Он показан на рис. 3 горизонтальной прямой. На рисунке также показан уровень значимости θ_0 , используемый для выбора модели. Абсциссы точек пересечения горизонтальной прямой $\theta = \theta_0$ с графиками зависимостей для величин $\theta_I(\eta)$ и $\theta_P(\eta)$ ($1/C_0$ и C_0) являются значениями границ величины η для выбора вариантов моделей. При $\eta \rightarrow 0$ от точки $1/C_0$, уровень значимости модели I_I возрастает от значения θ_0 , стремясь к 1 или 100%.

Возрастание уровня значимости модели I_P от значения θ_0 до 100% имеет место, когда $\eta \rightarrow \infty$ от точки C_0 .

Выбор варианта модели, в ряде случаев, можно осуществить, используя логику «мажорантных» соотношений. Суть состоит в следующем. Для заданного граничного уровня величины $\eta = C$, используя выражения (25) и (26), можно записать:

$$|I_P| = C|I_I| = C \left| \ln \frac{K_{e1}}{K_{e2}} \right| = C |\ln k_{ke}| = C \ln \frac{\max\{K_{e1}, K_{e2}\}}{\min\{K_{e1}, K_{e2}\}}. \quad (29)$$

Определим максимально возможное значение логарифма в правой части выражения (29). Известно, что максимальное значение величины энтропийного коэффициента $K_{e(max)} = 2.066 \approx 2.07$ имеет место в случае нормального закона распределения параметра [1–4, 29]. Теоре-

тически, значения величины K_e могут изменяться в пределах $0 \leq K_e \leq 2.07$. Однако, в реальных ситуациях, минимальное значение величины энтропийного коэффициента $K_{e(min)}=1$. Это наглядно показано в работах [1, 2]. Исходя из этого, представляется целесообразным ввести понятие практически предельной кратности энтропийного коэффициента или практически предельного значения коэффициента

преобразования закона распределения $k_{ke(max)} = \frac{K_{e(max)}}{K_{e(min)}} \approx 2.07$. Также,

для удобства формулировок, введем понятие кратности изменения величины СКО в виде $k_\sigma = \frac{\max\{\sigma_1, \sigma_2\}}{\min\{\sigma_1, \sigma_2\}}$. Теперь условие для оценки адекватности варианта информационной модели может быть сформулировано в виде следующей теоремы.

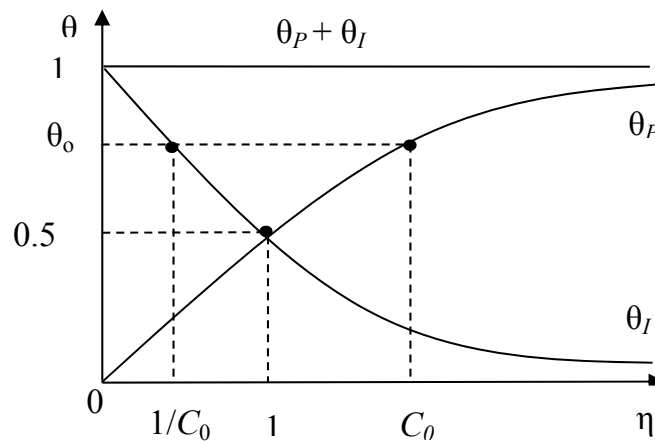


Рис. 3. Уровни значимости информационных моделей

Теорема. Если кратность изменения величины СКО на рассматриваемом этапе — k_σ превышает значение 2.07^C , то, со значением η , превышающим C , информационная модель может быть представлена величиной энергетической составляющей информации I_P .

Доказательство. Используя в (29) определение величины I_P , а также обозначения k_σ и значение величины $k_{ke(max)}$ можно записать:

$$|I_P| = \left| \ln \frac{\sigma_1}{\sigma_2} \right| = \ln \frac{\max\{\sigma_1, \sigma_2\}}{\min\{\sigma_1, \sigma_2\}} = \ln k_\sigma = C |\ln k_{ke}| \leq C \ln k_{ke(max)} = \ln 2.07^C. \quad (30)$$

Откуда следует, что при $\eta=C$, $k_{\sigma} \leq 2.07^C$. Если имеет место условие:

$$k_{\sigma} \succ 2.07^C, \quad (31)$$

то это означает, что значение величины $|I_P|$ увеличилось и, следовательно, согласно определению (26) имеет место соотношение:

$$\eta \succ C. \quad (32)$$

Что и требовалось доказать.

Используя эту теорему, представляется возможным осуществить обоснованный выбор варианта информационной модели (в данном случае в виде I_P). Исходными данными будут являться значение кратности изменения величин СКО на рассматриваемом этапе эволюции системы и значение граничного уровня коэффициента соотношения информации C . Уровень значимости такой модели, согласно (28), будет превышать величину $\theta_P = \frac{C}{1+C}$. На рис. 1 соответ-

ствующее значение величины θ_P будет находиться на участке кривой $\theta_P(\eta)$, расположенной по оси абсцисс правее точки C_0 . Получение информации о возможных значениях величины k_{σ} , характерных для конкретного этапа функционирования конкретной системы, может быть осуществлено различными способами: на основании экспериментальных или аналитических исследований, на основании исследований аналогов и др.

Для удобства использования и оперативности проведения вычислений условие (31) можно заменить более простым и компактным выражением:

$$k_{\sigma} \succ 2^C. \quad (33)$$

Относительная погрешность определения правой части неравенства (31) при переходе к упрощенному выражению (33) – γ будет определяться из выражения:

$$\gamma(\%) = \frac{2.07^C - 2^C}{2.07^C} \times 100(\%) \approx (1 - 0.97^C) \times 100(\%). \quad (34)$$

Так, например, согласно (34), для граничного уровня $C=3$, $\gamma \approx 9\%$, для $C=4$, $\gamma \approx 11\%$, для $C=5$, $\gamma \approx 14\%$ и т.д.

Полученные варианты информационных моделей также могут быть трансформированы и адаптированы к конкретной специфике их использования.

Нахождение энергетической составляющей информации, в соответствии с выражением (24), можно осуществить, используя известные методики определения величин СКО для линейных и нелинейных систем [5, 7, 9]. Это, например, возможно при наличии зависимости $x(\xi)$, описывающей эволюцию параметра x между рассматриваемыми этапами от какого-либо фактора ξ , и закона его распределения на начальном этапе – $p(x)$.

$$I_p = \ln \frac{\sigma_1}{\sigma_2} = \ln \sqrt{\frac{\int_{-\infty}^{\infty} (x - m_1)^2 p(x) dx}{\int_{-\infty}^{\infty} x^2(\xi) p(x) dx - m_2^2}} = \ln \sqrt{\frac{\int_{-\infty}^{\infty} (x - \int_{-\infty}^{\infty} x p(x) dx)^2 p(x) dx}{\int_{-\infty}^{\infty} x^2(\xi) p(x) dx - (\int_{-\infty}^{\infty} x(\xi) p(x) dx)^2}}. \quad (35)$$

В выражении (35) величинами m_1 и m_2 обозначены математические ожидания параметра в начале и конце этапа. Из выражения (8) видно, что величина, находящаяся под знаком логарифма, зависит только от характеристики $x(\xi)$ и закона распределения параметра на начальном этапе $p(x)$. По форме записи она фактически является величиной, обратной величине коэффициента статистической линеаризации для СКО. По сути, имеются существенные отличия. Коэффициент статистической линеаризации определяется на основе статической характеристики нелинейного звена. А характеристика $x(\xi)$ описывает эволюцию параметра системы в зависимости от значения какого-либо фактора, например, времени или пространственных координат. Поэтому, для удобства использования, представляется целесообразным определить выражения величин I_p для базовых стадий технологических процессов, реализуемых на типовом оборудовании, а также для реальных вариантов законов распределений параметров и их табулировать [12]. Таковыми, например, могут являться стадии прогрева, первичной и последующих видов тепловой обработки в автоклавах, печах для обжига изделий в металлургии, промышленности строительных материалов и др. Примерами этому являются тепловая обработка консервов в автоклаве, в соответствии с задаваемой формулой стерилизации, термообработка колбас и

др. [44, 45]. В этих процессах, изменения температурных режимов должны осуществляться по заданным временным зависимостям. В дальнейшем, использование готовых выражений для величины I_P позволит значительно упростить расчеты при проведении мониторинга и управления процессами, реализуемыми на таком оборудовании.

Выбор варианта модели может быть осуществлен не только аналитическим путем на основе коэффициента соотношения информации. В основу выбора могут быть положены иные соображения, базирующиеся, например, на анализе явлений, определяющих функционирование системы, на опытно-интуитивном восприятии ситуаций, на аналогиях с подобными системами и др.

Целесообразность использования информационных моделей подтверждена результатами исследований систем различной природы: экологических, технических, социальных и др. Эти результаты были использованы для прогнозирования эволюционных процессов, а также для выработки решений по модернизации систем и организации управления [1–5, 12, 44].

5.6. Приложение методов ТЭП к решению практических задач

Информационные портреты и портреты в пространстве характеристик энтропийных потенциалов обладают наглядностью, «выразительно» характеризуют изменение состояний неопределенности и могут быть использованы для проведения исследований и сравнительного анализа различных явлений, процессов и др. Образно говоря, они позволяют осуществить селекцию «вариативных свойств» состояний неопределенности.

Для иллюстрации изложенного на рис. 4 приведен качественный вид обобщенного фрагмента портрета в пространстве параметров энтропийных потенциалов σ и K_e (в виде отдельной траектории) состояния неопределенности температурного режима камеры для дефростации (размораживания) мясного сырья [5]. В качестве пояснения необходимо отметить следующее. Замороженное мясное сырье в виде полутуш свинины и говядины зачастую поступает на мясоперерабатывающие заводы из промышленных холодильников. Для последующей переработки, например, для нужд колбасного производства, сырье предварительно необходимо подвергнуть размораживанию. В противном случае бу-

дуг иметь место повышенный износ рабочих поверхностей обрабатывающего оборудования (режущих кромок ножей, поверхностей волчков и др.), а также повышенные затраты энергии при его работе.



Рис. 4. Фрагмент портрета в пространстве параметров энтропийных потенциалов состояния неопределенности температурного режима дефростационной камеры

Размораживание сырья осуществляется в специальных дефростационных камерах, в которых поддерживается определенный температурно-влажностный режим. Длительность процесса составляет сутки и более. Интенсивные изменения теплового режима приводят к значительным деформациям клеточных тканей и мембран, что, в конечном счете, может привести к их разрывам. В результате чего происходит вытекание «сока» и дефростированное сырье частично теряет вкусовые качества, питательную ценность и массу. Поэтому проблема контроля и управления состоянием неопределенности температурного поля в объеме камеры является актуальной.

Следует отметить, что подобные проблемы имеют место, например, в металлургической, стекольной промышленности при «расхолаживании» изделий или отливок заготовок для различных деталей. Так при остывании массы сваренного оптического стекла, интенсивные изменения теплового режима процесса приводят к появлению внутренних микротрещин и, в конечном счете, снижению его оптических свойств и др.

Приведенный на рис. 4 портрет строился на основании измерений значений температур в объеме дефростационной камеры. Измерения

проводились в течение суток через каждые три часа. Первый этап измерений совпадал с началом работы первой смены (8^{00}). (На предприятии существовал двухсменный режим работы.) Таким образом, было проведено восемь циклов контроля состояний неопределенности температурного поля, которым соответствуют точки 1–8 на рисунке. Каждой i -ой точке соответствуют свои значения величин σ_i и K_{ei} ($i=1,2,\dots,8$) на плоскости параметров ЭП. А каждое соответствующее значение величины ЭП, согласно (9), равно площади прямоугольника, образованного этими значениями этих параметров. Последовательность этапов, помимо возрастающей нумерации, также проиллюстрирована пунктирной кривой. Реальный диапазон изменения величин энтропийного коэффициента ($1 \leq K_e \leq 2.07$) обозначен соответствующими горизонтальными пунктирными линиями. На основании экспертных оценок, для отдельных производственных ситуаций выбиралось критическое значение величины ЭП $\Delta_e = K_e \sigma = C_{кр}$, превышение которого негативно сказывалось на качестве и свойствах дефростированного сырья. Этому условию соответствует множество точек, образующих «критическую» изотропу. Таким образом, все множество возможных вариантов состояний неопределенности температурного поля, соответствующих условию $\Delta_e \leq C_{кр}$, оказалось отображенным в область допустимых состояний неопределенности, которая обозначена на рис. 4 наклонной штриховкой. Из анализа представленного фрагмента энтропийного портрета можно сделать однозначный вывод о том, что с начала работы предприятия имело место нарастание состояния неопределенности температурного поля камеры. «Апогей» этого процесса имел место в «разгар» производственной деятельности. При этом на двух этапах наблюдений имел место выход из области допустимых состояний неопределенности (точки 4 и 5). Такая ситуация объясняется тем, что по мере нарастания темпов производственной деятельности также возрастало количество и интенсивность загрузок и выгрузок замороженного и дефростированного сырья. В результате частых открываний загрузочных дверей происходили процессы теплообмена с окружающей наружной средой, что и являлось одним из возмущений теплового режима. Кроме того, количество тепла, «вносимого» каждой отдельной массой партии загружаемого или выгружаемого сырья, также являлось возмущением теплового режима. По мере «сворачивания» производственной деятельности к концу второй смены также снижалась интенсивность перезагрузок

камеры, и состояние неопределенности температурного поля стало уменьшаться. В вечернее и ночное время, когда производство не работало, состояние неопределенности стабилизировалось на одном уровне (точки 7 и 8), весьма близком к исходному уровню (точка 1). На основании анализа траекторий энтропийных потенциалов были выработаны рекомендации по уменьшению влияния возмущений на состояние неопределенности температурного поля камеры. В качестве одного из решений было предложено установить дополнительный «шлюзовой» отсек перед входом в камеру. Реализация этой рекомендации позволила снизить уровень состояния неопределенности поля и, как следствие этого, изображающие точки в пространстве параметров энтропийных потенциалов «стянулись» ближе к исходной точке 1, войдя в зону допустимых состояний неопределенности.

В качестве дополнительной иллюстрации возможностей методов ТЭП по наглядному отображению эволюций состояний неопределенности на рис. 5 приведен информационный портрет выше рассмотренных состояний температурного поля дефростационной камеры.

Портрет построен по вышеизложенной методике, основанной на использовании выражений (23)–(25). В качестве исходной переменной функции I_L использовались номера этапов наблюдений – i , ($i=1,2,\dots,8$), что соответствует переходу к рассмотрению явления на временной решетке с периодом дискретизации равному трем часам. При этом предполагалось, что первый этап соответствует началу производственного процесса, которому предшествовало такое же установившееся состояние неопределенности температурного поля при $i=0$ (по аналогии с этапами 7 и 8), то есть $\Delta_{e1}=\Delta_{e0}$. Отсюда и следует, что $I_{L(0-1)}=0$. В примере, приведенном на рис. 4, $K_{e1}\approx K_{e2}$, и, следовательно, на этом этапе, $I_I = \ln \frac{K_{e1}}{K_{e2}} \approx 0$. Поэтому согласно (24) $I_{L(1-2)} \approx I_P = \ln \frac{\sigma_1}{\sigma_2}$.

То есть величину информации, порожденной изменением состояния неопределенности температуры на этом этапе, можно определять по величине энергетической составляющей. Аналогичная ситуация имеет место на этапах (5–6) и (6–7). В этих случаях для вычисления величин $I_{L(5-6)}$ и $I_{L(6-7)}$ также можно использовать выражение (24). А на этапе (4–5) имеет место ситуация: $\sigma_4 \approx \sigma_5$, следовательно $I_P = \ln \frac{\sigma_4}{\sigma_5} = 0$.

Поэтому согласно (25) $I_{L(4-5)} \approx I_I = \ln k_{ke(4-5)} = \ln \frac{K_{e4}}{K_{e5}}$. Другими словами,

величину информации, порожденной изменением состояния неопределенности температуры на этом этапе, можно определять по величине интеллектуальной составляющей. Ситуации на этапах (2–3) и (3–4) носят «общий» характер, где ни одна из составляющих информации не равна нулю, поэтому для вычисления соответствующих значений величин $I_{L(2-3)}$ и $I_{L(3-4)}$ необходимо использовать выражение (23).

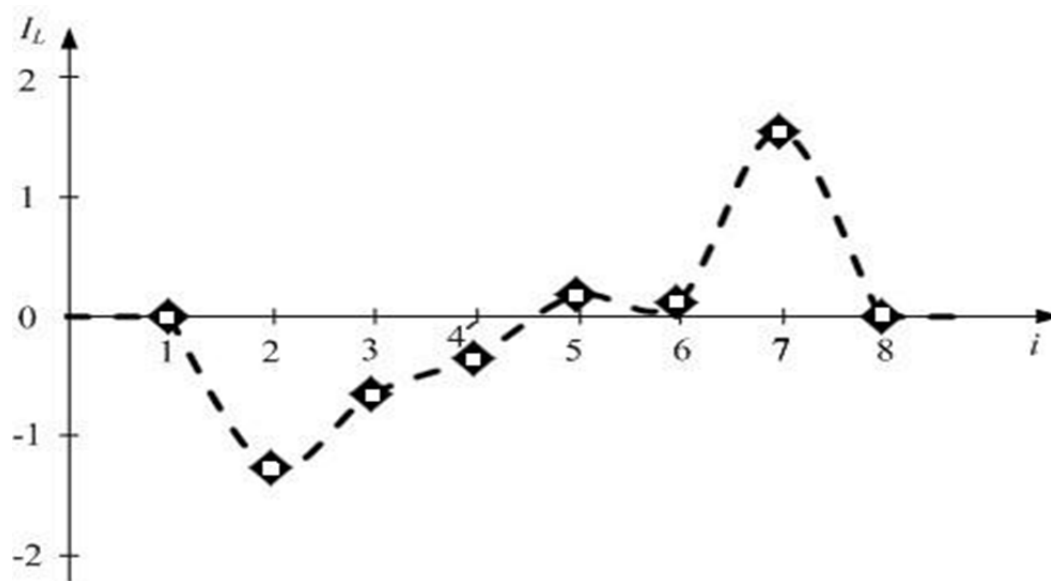


Рис. 5. Информационный портрет состояний неопределенности температурного поля дефростационной камеры

Информационный портрет в иллюстративном плане «дополняет» возможности энтропийного портрета и не противоречит ему. Как было отмечено выше, он позволяет наглядно показать тенденцию или тренд в изменении состояний неопределенности. Как показано на рис. 5 наибольшие изменения этих состояний имели место при переходе к этапам 2 и 7. В первом случае имело место нарастание состояния неопределенности ($I_{L(1-2)} < 0$), а во втором – наоборот ($I_{L(6-7)} > 0$).

Приведенные примеры (рис. 4 и рис. 5) являются наглядной иллюстрацией возможностей информационных и энтропийных портретов для описания и исследования состояний неопределенности различных параметров. И в этом смысле их можно рассматривать как элементы когнитивных изображений этих состояний.

Представляется целесообразным отметить сферы или области, где применение предложенного подхода оказалось продуктивным или имеет перспективы дальнейшего развития и внедрения.

1. Организация мониторинга и управления различными техническими объектами: динамическими системами, теплообменными аппаратами, устройствами и технологическими процессами для производства продукции в пищевой промышленности, дефростационными холодильными и климатическими камерами [9–12, 15, 46–48] и др.

2. Организация мониторинга и прогнозирования «поведения» сложных систем и явлений: сейсмических процессов [16], климатических изменений [1, 3], электрокардиографических (ЭКГ) сигналов [36], инфраструктурных преобразований и финансовой деятельности [49, 50]. Особенностью функционирования такого рода систем является возможность появления различных видов случайных составляющих в протекающих процессах. Для «стороннего» наблюдателя каждое конкретное значение параметра может рассматриваться как реализация случайной величины, имеющей соответствующую вероятность появления. Поэтому вариации анализируемого параметра в таких процессах могут быть охарактеризованы уровнем его состояния неопределенности в течение какого-либо периода. Указанное обстоятельство создает основу для применения методов ТЭП к проведению соответствующих исследований. Исследование вариативных свойств параметра на основе понятий энтропийных потенциалов позволяет осуществлять мониторинг состояний таких процессов, выявлять намечающиеся тенденции в их развитии.

3. Обработка результатов наблюдений [1, 13]. Особенности приложения методов ТЭП к обработке наблюдений состоят в следующем. Любые наблюдения содержат случайные составляющие, ошибки, которые порождают состояние неопределенности «знаний» (например, в виде каких-либо моделей), извлекаемых из результатов наблюдений. Поэтому «качество» получаемой математической модели, может быть охарактеризовано уровнем ее состояния неопределенности по описанию исследуемого параметра. Отсюда следует, что определение параметров модели, необходимо осуществлять путем минимизации какого-либо критерия, характеризующего ее состояние неопределенности. В качестве такового целесообразно использовать какой-либо из вариантов критериев на основе набора понятий энтропийных потенциалов, а также информационные критерии. Особенности выбора конкретных вариантов рассмотрены в вышеуказанных работах.

Заключение

ТЭП и разработанные на ее основе методы позволяют получать количественные оценки состояний неопределенности различных систем и изменения этих состояний в виде величин Δ_e , L_Δ , La_z , I_L . Эти величины, основанные на соответствующих понятиях, являются не противоречивыми и основаны на принципах «вложения»: величины более высокого уровня выражаются через величины предыдущих уровней путем учета дополнительных характеристик параметров системы и наоборот. Набор понятий величин энтропийных потенциалов позволяет описать разные «уровни» состояний неопределенности в зависимости от специфики поставленной задачи. Величины энтропийных потенциалов L_Δ и La_z являются безразмерными, что позволяет использовать их в качестве критериев подобия систем различной природы при исследовании состояний неопределенности на основе «энтропийного подхода». Тем самым создаются предпосылки для проведения такого рода исследований методами теории подобия. Кроме того, наличие вышеуказанных количественных оценок состояний неопределенности позволяет определять устойчивость этих состояний, в различных ситуациях, используя классические подходы и определения. Также с использованием предложенных понятий оказывается возможным проводить сравнительный анализ функционирования различного оборудования, протекания различных процессов и явлений, особенно в условиях неопределенности.

Особенностью и достоинством данного подхода является удобство, и наглядность получаемых результатов, а также простота их интерпретации. Он основан на комплексном использовании наглядных информативных характеристик K_e , X_n , σ . Эти характеристики поддаются достаточно простому определению. В случаях, когда они определяются на основании результатов наблюдений, для их определения требуется меньший объем данных, чем, например, для определения значений величин энтропий параметров. Получаемые модели состояний неопределенности оказываются достаточно адекватными существующим реалиям, компактными и удобными для практического применения.

Разработанные методы отображений указанных характеристик в величины введенных понятий энтропийных потенциалов, I_L и др. обоснованы, а их свойства и возможности исследованы в рамках ТЭП. Результаты разработок и исследований оформлены наборами определений, лемм, теорем и соответствующих иллюстраций.

Использование введенных понятий позволяет осуществить наглядную графическую иллюстрацию изменения состояний неопределенности систем в виде информационных и энтропийных портретов различной размерности. Эти портреты позволяют наглядно представлять и исследовать эволюцию этих состояний. Их также можно рассматривать, как элементы когнитивной графики, использование которых позволяет осуществить «сжатие» исходной информации о свойствах и состояниях систем.

Методы ТЭП могут быть применены к исследованию систем любой природы, с вероятностными и детерминированными связями, так как проявление детерминированных связей, при изменении рамок и размерности задачи, также могут рассматриваться как проявление состояний неопределенности.

Список литературы

1. *Лазарев В.Л.* Теория энтропийных потенциалов: Монография. С-Пб.: Изд-во Политехнического ун-та. 2012. 127 с.
2. *Lazarev V.L.* The Theory of Entropy Potentials, Basic Concepts, Results and Applications // Pattern Recognition and Image Analysis. (Advances in Mathematical Theory and Application.) 2011. Vol. 21. № 4. Pp. 637–648.
3. *Lazarev V.L.* Analysis of Systems Based on Entropy and Information Characteristics. Technical Physics, 2010. Vol. 55. № 2. Pp. 159–165.
4. *Lazarev V.L.* An entropy approach to monitoring and control. Journal of Computer and Systems Sciences International. 2005. Vol. 44. № 6. Pp. 893–899.
5. Управление в условиях неопределенности: Монография / С.В. Прокопчина, М.Ю. Шестопапов, Л.В. Уткин, М.С. Куприянов, В.Л. Лазарев, Д.Х. Имаев, В.Л. Горохов, Ю.А. Жук, А.В. Спесивцев. Под ред. С.В. Прокопчиной. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2014. 304 с.
6. *Лазарев В.Л.* Информационные модели эволюций систем. Сб. докл. XIII Междунар. научно-практич. конференции «Теория и практика современной науки», г. Москва, 2–3 апреля 2014 г. Т. 1. Изд-во «Спецкнига», 2014. 452 с. С. 200–204.
7. *Лазарев В.Л.* Практические аспекты исследования систем в информационном пространстве / Сб. докладов XVII Международной конференции по мягким вычислениям и измерениям SCM'2014. С-Пб., 21-23 мая 2014 г. Изд-во СПбГЭТУ «ЛЭТИ». 2014. Т. 1. С. 42–45.

8. *Lazarev V.L.* Representative information models for monitoring and control in the conditions of uncertainty // XVIII Intern. Confer. «SCM». 2015. IEEE. Pp. 54–57. DOI: 10.1109/SCM.2015.7190408.

9. *Лазарев В.Л., Митин Е.Е.* Мониторинг процессов термообработки жидких продуктов на основе методов теории энтропийных потенциалов // Вестник Международной академии холода. 2013. № 2. С. 43–45.

10. *Лазарев В.Л., Иванов В.Л., Фролков Н.А.* Когнитивные информационные технологии мониторинга объектов пищевой промышленности на основе энтропийных потенциалов параметров // Вестник Международной академии холода. 2015. № 2. С. 66–70.

11. *Лазарев В.Л., Грахольская Т.А., Травина Е.А., Фролков Н.А.* Использование когнитивных образов состояний систем в пространстве параметров энтропийных потенциалов для организации мониторинга и управления // Научный журнал НИУ ИТМО. Серия «Процессы и аппараты пищевых производств». 2015. № 4. С. 54–61.

12. *Лазарев В.Л.* Совершенствование управления с использованием характеристик энтропийных потенциалов. Адаптация к специфике биотехнологической промышленности // Вестник Международной академии холода. 2016. № 4. С. 68–73.

13. *Lazarev V.L.* Processing of observations on the basis of information criteria // Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016. Saint Petersburg, RUSSIA. Publisher: IEEE. Pp. 48–50. DOI: 10.1109/SCM.2016.7519679.

14. *Lazarev V.L.* Epistemological Foundations for Generation of Perspective Competencies in the Training of Personnel for Industrial and Economic Complex. 4th Forum Strategic Partnership of Universities and Enterprises of Hi-Tech Branches (Science, Education, Innovations). Saint Petersburg, RUSSIA. NOV 11–13. Pp. 26–28. Publisher: IEEE. DOI: 10.1109/IV Forum.2015.7388242.

15. *Лазарев В.Л.* Исследование и проектирование многокомпонентных систем в биотехнологической промышленности в условиях неопределенности на основе энтропийных потенциалов показателей качества // Вестник Международной академии холода. 2017. № 2.

16. *Лазарев В.Л., Митин Е.Е.* Мониторинг сейсмических процессов на основе энтропийных потенциалов сейсмограмм // Проблемы современной науки и образования – 2016. – № 16(58). – С. 139–147.

17. *Лазарев В.Л., Иванов В.Л., Фролков Н.А.* Исследование эффективности функционирования системы управления котлоагрегатом на основе вероятностно-энтропийного подхода // Международная кон-

ференция по мягким вычислениям и измерениям – 2015. – Т. 1. – № Секции 1–3. – С. 149–152.

18. Лазарев В.Л., Иванов В.Л., Фролков Н.А. Перспективные направления организации робастного управления в биотехнологической промышленности // VII Международная научно-техническая конференция «Низкотемпературные и пищевые технологии в XXI веке» (Санкт-Петербург, 17–20 ноября 2015 г.): Материалы конференции – 2015. – № 1. – С. 279–282.

19. Lazarev V.L., Ivanov V.L. Product quality monitoring in industrial biotechnology within the space of entropy potentials parameters // Proceedings of the 19th International Conference on Soft Computing and Measurements, SCM 2016. 2016. Pp. 218–220.

20. Интеллектуальные системы автоматического управления / под ред. И.М. Макарова, В.М. Лохина. М.: Физматлит, 2001. 576 с.

21. Охмилев М.Ю., Соколов Б.В., Юсупов Р.М. Интеллектуальные технологии мониторинга и управления структурной динамикой сложных технических объектов. М.: Наука, 2006. 410 с.

22. Zadeh L. Fuzzy logic, neural networks and soft computing // Commutation on the ASM. 1994. Vol. 37. № 3. Pp. 77–84.

23. Уткин Л.В. Анализ риска и принятие решений при неполной информации. С-Пб.: Наука, 2007. 404 с.

24. Прокопчина С.В. Основные принципы решения задач измерения и управления сложными антропогенными объектами в условиях неопределенности / Сб. докладов Международной конференции по мягким вычислениям и измерениям «SCM-2012». В 2-х томах, С-Пб. (Пленарный доклад.) Изд-во СПбГЭТУ «ЛЭТИ», 2012. Т.1. С. 23–34.

25. Горохов В.Л., Муравьев И.П. Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Методология, методики и интерфейсы. С-Пб.: ИНЖЕКОН, 2007. 173 с.

26. Тулупьев А.Л., Николенко С.И., Сироткин А.В. Байесовские сети: Логико-вероятностный подход. С-Пб.: Наука, 2006. 607 с.

27. Методы классической и современной теории управления. Учебник: в 5 т. / под ред. Н.Д. Егупкина. М.: МГТУ им. Н.Э. Баумана, 2005. Т.1. Анализ и статистическая динамика систем автоматического управления. 748 с. Т.2. Синтез регуляторов и теория оптимальных систем автоматического управления. 736 с. Т.3. Методы современной теории автоматического управления. 748 с.

28. Дюбуа Д., Прад А. Теория возможностей. М.: Радио и связь, 1990. 290 с.

29. Электрические измерения неэлектрических величин. А.М. Турчин, П.В. Новицкий, Е.С. Левшина и др. / под ред. П.В. Новицкого. Л.: «Энергия», 1975. 576 с.

30. Лазарев В.Л. Технологии моделирования и управления состояниями неопределенности систем на основе теории энтропийных потенциалов / Сб. докладов 5-й Всероссийской научной конференции «Управление и информационные технологии» (УИТ-2008). В 2-х томах. С-Пб.: Изд-во СПбГЭТУ «ЛЭТИ», 2008. Т. 2. С. 10–14.

31. Патент № 2296356 RU, МПК G05B13/00. Способ контроля и управления динамической системой / Лазарев В.Л. (RU), С 1, 2007.03.27.

32. Патент № 2334262 RU, МПК G05B13/00. Способ контроля и управления состоянием неопределенности системы / Лазарев В.Л. (RU), С 1, 2008.09.20.

33. Кулаков В.Г., Лазарев В.Л., Федулин В.А. Энтропийные модели в исследовании социальных систем // Вопросы статистики. 2010. № 10. С. 47–50.

34. Лазарев В.Л., Федулин В.А. Мониторинг и организация устойчивого развития территорий на базе теории энтропийных потенциалов / Сб. докладов VII Международной научно-практической конференции «Природноресурсный потенциал, экология и устойчивое развитие регионов России». Пенза: МНИЦ ПГСХА, 2009. С. 83–86.

35. Каминская В.И. Управление процессами намыва сооружений в условиях неопределенности // Гидротехническое строительство. 2009. № 1. С. 27–29.

36. Лазарев В.Л., Митин Е.Е. Исследование электрокардиографических сигналов с использованием методов теории энтропийных потенциалов / Сб. докладов Международной конференции по мягким вычислениям и измерениям SCM'2012. В 2-х томах. С-Пб.: Изд-во СПбГЭТУ «ЛЭТИ», 2012. Т. 2. С. 123–126.

37. Лазарев В.Л. Мониторинг сложных систем в экологии, экономике, социальной сфере методами теории энтропийных потенциалов / Сб. докладов XXII Международной научно-технической конференции «Математические методы и информационные технологии в экономике, социологии и образовании». Пенза: Приволжский Дом знаний. 2008. С. 197–200.

38. Лазарев В.Л. Организация процессов нанотехнологий на основе энтропийных потенциалов параметров / Сб. докладов Междуна-

родной конференции «Низкотемпературные и пищевые технологии в XXI веке». С-Пб., 2007. С. 78–85.

39. Фазовое пространство. Математическая энциклопедия / гл. редактор И.М. Виноградов: Изд-во «Советская энциклопедия». М., 1985. Т. 5. С. 587–588.

40. *Месяц Г.А.* Импульсная энергетика и электроника. М.: Наука, 2004. 704 с.

41. *Собельман И.И.* Введение в теорию атомных спектров. М.: Наука, 1977. 320 с.

42. *Борн М., Вольф Э.* Основы оптики. М.: Наука, 1970. 855 с.

43. *Кондраков Г.В., Рязанова М.П.* Общая теория материи. Изд. 2-е, перераб. и доп. С-Пб.: ХИМИЗДАТ, 2008. 592 с.

44. *Лазарев В.Л.* Робастное управление в биотехнологической промышленности. Учебное пособие. СПб.: Университет ИТМО, 2015. 196 с.

45. *Лазарев В.Л., Иванов В.Л.* Системный подход к оценке состояний неопределенности. Особенности реализации для задач мониторинга и управления // Вестник Международной академии холода. 2014. № 2. С. 25–29.

46. *Нечаев Ю.И., Петров О.Н.* Функциональное пространство при интерпретации «потенциальной ямы» в условиях резкого снижения остойчивости поврежденного судна // Морские интеллектуальные технологии. 2014. № 26. С. 15–21.

47. *Бураковский П.Е., Нечаев Ю.И.* Практическая реализация графоаналитической системы контроля общей прочности промысловых судов на основе современной теории катастроф // Морские интеллектуальные технологии. 2014. № 4(24). С. 10–16.

48. *Краснов А.Е., Красуля О.В., Большаков О.В., Шленская Т.В.* Информационные технологии пищевых производств в условиях неопределенности. (Системный анализ управления и прогнозирования с элементами компьютерного моделирования). М.: ВНИИМ, 2001. 496 с.

49. *Домбровский М.А., Буторина О.В.* Анализ энтропийного потенциала инфраструктуры региона на примере Пермского края // Экономика и предпринимательство. 2015. № 5-1 (58-1). С. 305–310.

50. *Кошевой О.С., Глебова Т.А., Чиркина М.А.* Диагностика показателей концентрации распределения среднедушевых денежных доходов населения на основе энтропийного подхода // Региональная архитектура и строительство. 2012. № 2. С. 192–195.

ГЛАВА 6

МЕТОДОЛОГИЯ КОГНИТИВНОЙ ВИЗУАЛИЗАЦИИ МНОГОМЕРНЫХ ДАННЫХ

В.Л. Горохов, Ю.В. Барышев, В.В. Витковский

Введение

Когнитивные технологии как основа для нового поколения систем поддержки интеллектуальных решений

Разработка современных сетевых технологий для развития электронного бизнеса и формирования структур «электронного правительства» в рамках социально-экономической системы России требует инновационных путей развития не только в области информационных технологий, но и инноваций в социокультурной сфере. Эти инновации требуют более глубокого учета человеческого фактора. Подобный учет «человеческого капитала» становится возможен на основе фундаментальных достижений когнитивных наук. При этом техническим инструментом для учета человеческих интенций, интуиции будут выступать программные системы типа интеллектуальных интерфейсов. Ключевую методическую роль при разработке таких систем и самое главное при их эксплуатации играет современная когнитивная психология и теория познания (гносеология). Без исследования фундаментальных аспектов формирования когнитивных образов в системах интеллектуальных интерфейсов невозможно системно реализовать сценарий инновационного развития системы электронного бизнеса и систем поддержки «электронного правительства» России. Таким образом, можно сформулировать важное направление научных исследований, на развитие которых требуется направить серьезные усилия. Это исследование особенностей и разработка методических элементов анализа систем когнитивной машинной графики, в контексте инновационного сценария для развития интеллектуальных систем поддержки принятия решений и систем интеллектуальных агентов.

6.1. Когнитивные технологии как основа для нового поколения систем поддержки интеллектуальных решений

6.1.1. Современное состояние исследований по данной проблеме, основные направления исследований в мировой науке

Термин когнитивная машинная графика введен в научный оборот известным математиком В.С. Зенкиным получил широкое распространение в западной и отечественной науке. Учитывая многоплановость проблем формирования и развития когнитивных технологий, их взаимосвязь с различными аспектами информатики и экономики, отдельные вопросы данной тематики изучались многими отечественными и зарубежными учеными самых различных школ и направлений (Данной проблематике посвящены труды ученых Allen R.B. Allport D.A., Buja D., Cook D., F. Swayne, Koschat M., Swayne D.F. Symanzik J., Majure B., Cook D. and Cressie N., Geladi P, Esbensen K, Прокопчина С.В., Тарасов В.Б., Аверкин А.Н., Лукьянец А.А., Горохов В.Л, Ротарь В.Г., Чернов А.Г. Кулинич А.А.). Вместе с тем, до сих пор остается множество вопросов, требующих дальнейшей разработки.

В имеющихся работах перечисленных авторов значительное внимание уделяется таким проблемам, как выявление и генерация самих когнитивных эффектов. Вместе с тем, понятийная структура и терминология в данной области философских, психологических, информационных и экономических исследований еще не устоялась и требует уточнения и систематизации. По-прежнему недостаточно раскрыты основные черты и характеристики механизмов генерации когнитивных образов для стимуляции интуиции и генерации новых знаний как системного и философского феномена, остаются не раскрытыми роль отдельных алгоритмов и психологических феноменов в формировании когнитивных образов и ряд других проблемных областей когнитивных наук.

Можно констатировать, что в самом общем виде теория генерации когнитивных образов для стимуляции интуиции и формирования принципиально новых знаний как целостная концепция, описывающая реальные изменения в современной когнитивной науке и гносеологии, находится в процессе становления и требует комплексного изучения и научного анализа данной тематики.

6.1.2. Направление исследований

Исследование в данной области носит комплексный характер и тема может быть раскрыта следующими направлениями работ: анали-

тический обзор зарубежного и отечественного опыта исследования процессов развития когнитивных информационных систем, OLAP-технологий и визуального программирования в контексте инновационного развития сетевых технологий; определение и анализ основных тенденций и направлений инновационного развития электронного бизнеса и стратегий развития «электронного правительства» России; разработка методов и методик работы с когнитивными образами в системах когнитивной машинной графики; оценка влияния внедрения когнитивных инновационных технологий на развитие корпоративных информационных систем инновационной экономики; выявление сущности и роли системы формирования когнитивных образов в процессах стимуляции профессиональной интуиции лиц, принимающих решение в рамках инновационного развития на основе концепции человеческого капитала; разработка системы ключевых показателей состояния и развития систем когнитивной машинной графики в условиях инновационного развития; оценка основных параметров соответствия структуры системы интеллектуальной поддержки принятия решений для развития сетевых технологий электронного бизнеса.

Ниже рассмотрены некоторые элементы проводимых исследований, которые отражают новизну следующих направлений: классификация основных факторов, определяющих эффективность процесса формирования когнитивных образов в системах поддержки интеллектуальных решений в современном электронном бизнесе инновационной экономики; закономерности развития системы формирования когнитивных образов как элемента концепции человеческого капитала; состав и описание структуры когнитивных образов, критических элементов и процессов, влияющих на стимуляцию интуиции пользователя; результаты сравнительного анализа различных моделей построения когнитивных систем; классификация механизмов генерации когнитивных образов, обеспечивающих стимуляцию и интуицию для принятия решений в задачах инновационного развития экономики; спецификация ключевых показателей систем когнитивной машинной графики; методы оценки основных параметров соответствия систем когнитивной машинной графики потребностям основных инновационных кластеров экономики.

Исследование базируется на использовании концепций когнитивной психологии, концепций гносеологии, включающей идеи Э. Гуссерля и А. Бергсона, методах философского анализа на основе современных достижений онтологии (работы Дэвида Чалмерса и Дениела Денета). Предметом исследования является фундаментальные алгебраиче-

ские и топологические свойства когнитивных образов, философские концепции метафорического мышления в свете освоения человечеством новых мультимедиа технологий, тенденции и перспективы изменения и развития когнитивных информационных систем для задач инновационного развития; способы и методы философского описания и прогнозирования этих систем и когнитивных процессов.

Ниже рассмотрены первые примеры закономерностей формирования когнитивных образов, затем разработаны элементы методологии манипуляции этих образов с целью усиления эффектов суггестии. Исследование специализируется в части применения когнитивных систем для задач анализа многомерных данных природного и техно-сферного мониторинга, проблематики эффективного использования ресурсов интуиции человека в рамках теории человеческого капитала.

При выполнении исследования учитываются наработки в области когнитивистики и гносеологии; отечественные и зарубежные публикации и материалы по проблемам развития когнитивных наук, по теоретическим и методическим аспектам исследования человеческого и интеллектуального капитала, экономики знаний, интеграции сетевых технологий и бизнеса.

Приводимые ниже результаты опираются на ряд выполненных научно-исследовательских и прикладных работ, в том числе по проблематике формирования когнитивных образов, по мониторингу и анализу закономерностей развития сферы сетевых интеллектуальных интерфейсов, проблемам развития сетевых технологий, формированию принципов когнитивного визуального программирования и анализа многомерных данных с целью «добывания знаний» (DATA MINING) для инновационного развития систем научного, образовательного, экологического, медицинского и промышленного мониторинга Санкт-Петербурга. Эти результаты позволяют надеяться на успех дальнейших исследований в данной области знаний.

Прежде всего, следует отметить разработку нового поколения сетевых версий программ визуализации многомерных данных основанных на концепциях и научных метафорах гносеологии Э. Гуссерля. Здесь удалось разработать методы генерации псевдо – шестимерных когнитивных образов, которые успешно отражают ранее недоступные человеческой интуиции свойства многомерных данных.

Разработаны методики манипуляции с когнитивными образами обеспечивающие получение новых научных результатов в ряде отраслей фундаментального естествознания и геомониторинга.

Показана эффективность применения когнитивных информационных систем в области логистики, обработки астрофизических данных, медицинской статистики и мониторинга образовательных процессов в России.

6.1.3. Список основных публикаций

1. Д.В. Минаев, А.Б. Матюхин, В.А. Евдокимов. Проект разработки интерактивной персональной системы рейтингования вузов // «Менеджмент XXI века: образование и бизнес»: VII Международная научно-практическая конференция 29–30 ноября 2007 года: Тез. Докл. СПб.: РГПУ. 2007.

2. Д.В. Минаев, А.Б. Матюхин, В.А. Евдокимов. Опыт создания и функционирования интерактивной персональной системы рейтингования вузов на базе портала СПбГИЭУ (ИНЖЭКОН) // «Планирование и обеспечение подготовки кадров для промышленно-экономического комплекса региона»: VI Всероссийская научно-практическая конференция 6-7 декабря 2007 года.: Тез. докл. СПб.: СПбГЭУ «ЛЭТИ», 2007.

3. Gorokhov V., Vitkovskiy V., Komarinskiy S., Zakhrevski D. Space Walker. The Cognitive Visualization System With the Dynamic Projection of Multi-dimensional Data / Proceedings International conference «Astronomical Data Analysis Software and systems». London UK 23-26 September 2007. P. 51.

4. Vitkovskiy V., Gorokhov V. The Methods of Cognitive Visualization for The Astronomical Databases Analyzing Tools Development Space Walker. The cognitive visualization system with the dynamic projection of multidimensional data / Proceedings International conference «Astronomical Data Analysis Software and systems». London UK 23-26 September 2007. P. 51–52.

5. Gorokhov V. The Cognitive Visualization System the Dynamic Projection / Proceedings International conference «Astronomical Data Analysis Software and systems». London UK 23-26 September 2007. P. 36. Vitkovskiy V.V., Gorokhov V.L. New method of data mining in practical cosmology / Proceedings International conference «Problems of Practical Cosmology» Saint-Petersburg 2008. P. 256–259. (http://ppc08.astro.spbu.ru/fin_program.html).

6. Gorokhov V.L, Evdokimov V.A, Vitkovskiy V.V. Cognitive multidimensional data visualization in analyzing and decision_support systems / Третья Международная конференция по когнитивной науке 20–25 июня 2008 г., Москва, Россия, Тезисы докладов том 1 (The Third international conference on cognitive science june 20–25, 2008, Moscow, Russia abstracts volume 1). Москва, 2008. С. 55.

7. Струков Д.Р., Горохов В.Л. Когнитивные методы многомерной биомедицинской статистики для геоинформационных систем / Тезисы 10 Международной конференции «Региональная информатика 2008», РИ-2008. Санкт-Петербург 22–24 октября, 2008.

8. Komarinskiy, S.; Vitkovskiy, V.; Gorohov, V.; Zakharevski, D. The Cognitive Visualization System Astronomical Data Analysis Software and Systems (ADASS) XVII / Editors: Argyle, Robert W.; Bunclark, Peter S.; Lewis, James R. London, UK 2008. 394 p. (коллективная монография).

9. Лукьянец А.А., Горохов В.Л., Ротарь В.Г., Шумский А.А. Методы оценки доступности коммунальных услуг для населения Вестник ИНЖЕКОНА. Серия Экономика. Выпуск 1 (20). 2008. С. 77–84.

10. Горохов В.Л., Барышев Ю.В., Тееррикорпи П., Витковский В.В. Интеллектуальные адаптивные интерфейсы Data Mining для наблюдательной космологии / Сборник научных трудов научно-практической конференции по современным проблемам прикладной информатики. 27–28 мая 2008 г. СПб.: Издательство СПбГПУ, 2008. С. 158–166.

11. Горохов В.Л., Гусева Е.А., Леонов В.Е. Смысл понятия «информация» в теории постиндустриального общества / Материалы международной конференции «Глобальный мир: гуманитарный кризис или момент развития» апрель 2008 г. СПб.: Издательство СПбГПУ, 2008. С. 264–269.

12. Горохов В.Л., Гусева Е.А., Леонов В.Е. Актуальные проблемы современного естествознания / Тезисы докладов V научно-практической конференции «Отечественные традиции гуманитарного знания: история и современность СПбГИЭУ 22 мая 2009 г. СПб.: Издательство СПбГИЭУ, 2009. С. 28–31.

13. Горохов В.Л., Тееррикорпи П., Барышев Ю.В., Муравьев И.П. Компьютерные метафоры феноменологических конструкций Ф. Брентано, М. Хайдегера, Ж.П. Сартра для систем обработки астрофизических наблюдений / Труды международной конференции по мягким вычислениям, 25–27 июня 2009, (SCM – 2009). Санкт-Петербург, 2009.

14. Горохов В.Л., Ю.В. Барышев, П. Тееррикорпи, В.В. Витковский. Когнитивная визуализация космологических обзоров без привязки к сетке космологических моделей / Сборник научных трудов научно-практической конференции по современным проблемам прикладной информатики» 19–20 мая 2009 г. СПб.: Издательство СПбГПУ, 2009. С. 146–149.

15. Струков Д.Р., Горохов В.Л. Геоинформационные системы и многомерные статистические методы пространственного анализа для

исследования заболеваемости // Информационно-управляющие системы. Т. 3(40). 2009. С. 56–62.

16. *Gorokhov V.* Cognitive imaginig in visual data-driven decision-support systems / Proceedings International conference «Astronomical Data Analysis Software and systems» Sapporo. Japan. 4–8 October 2009.

17. *Evdocimov V, Gorokhov V., Vitkovskiy V.* Cognitive imaginig in visual data / Abstract. Scientific programme. International conference «High Performance Computing in Observational Astronomy: Requirements and Challenges». Puna. India. 12–14 October. 2009. P. 5.

По материалам исследований опубликованы монографии:

1. *Горохов В.Л., Муравьев И.П.* Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Методология, методики и интерфейсы. Монография. СПб.: ИНЖЭКОН, 2007. 173 с.

2. *Горохов В.Л., Лукьянец А.А., Чернов А.Г.* Современные методы когнитивной визуализации многомерных данных. Монография. Томск.: Некоммерческий фонд развития региональной энергетики, 2007. 216 с.

3. *Komarinskiy, S.; Vitkovskiy, V.; Gorohov, V.; Zakharevski, D.* The Cognitive Visualization System Astronomical Data Analysis Software and Systems (ADASS) XVII / Editors: Argyle, Robert W.; Bunclark.

4. *Горохов В.Л.* Управление в условиях неопределенности / Прокочина С.В., Шестопапов М.Ю., Уткин Л.В., Куприянов М.С., Лазарев В.Л., Имаев Д.Х., Горохов В.Л., Жук Ю.А., Спесивцев А.В. / Монография. СПб.: Из-во СПбГЭТУ «ЛЭТИ», 2014. 303 с.

5. *Горохов В.Л.* Быть или не быть: проблема мироздания в современной науке и технике / Семенов А., Малышев Ю., Горохов В. / Монография. Palmarium Academic Publishing. Ist ein Imprint der / OmniScriptun GmbH & Co. KG Heinrich-Bocking-Str. 6-8, 66121 Saarbrücken, Deutchland 2015. P. 97.

6.2. Расширение и обновление поля рационального анализа данных на основе когнитивных интерфейсов

Современное бурное развитие NBIC технологий радикально изменяет стратегию взаимодействия человека и компьютера. Огромные возможности, открытые в рамках когнитивных технологий [1–4] по-

казали принципиально новые аспекты и перспективы информатики в плане управляемого воздействия на познавательные возможности человека оператора. Это нашло свое воплощение в создании новых классов программных и аппаратных средств получивших название интеллектуального интерфейса и когнитивной компьютерной графики [4]. Общение человека и компьютера требует создания таких программных средств, которые обеспечат симметричную среду общения человека и компьютера. Речь идет о том, что огромные ресурсы компьютера могут быть более эффективно использованы, если программные системы и алгоритмические средства, на которых они построены, будут адекватно использоваться человеком – оператором.

Прежде всего, для удобства и правильного использования человеком компьютера необходимо создать средства, которые в диалоге с человеком определяли бы уровень общения, на котором человек оператор желает и может «общаться» с компьютером. Это позволит избежать недоразумений и конфликтов между пользователем и программной системой, между пользователями разных категорий. Такие средства и получили название интеллектуальных интерфейсов [4–5]. Подобные средства, обладая универсальностью и возможностью формировать диалог в любой предметной области (разумеется, после этапа обучения), могут обеспечить адекватный для пользователя диалог. Таким образом, может быть преодолена существующая ныне проблема «плохих» интерфейсов управления компьютерами (проблема ЮЗЕБИЛИТИ). Теоретической основой для создания таких интеллектуальных интерфейсов «сотрудничающих» с человеком является теория категорий.

Не менее важным для создания подобных интеллектуальных интерфейсов является неожиданный прорыв в области когнитивных технологий и гносеологии. Речь идет об экспериментальных открытиях (в рамках технологий искусственного интеллекта) феноменов сознания, которые позволяют актуализировать и активировать многие неиспользуемые в общении с компьютером возможности человека (профессиональная интуиция, эмпирические навыки и т.д.). Прежде всего, это установление прямой тождественности фреймов М. Минского и нозм Э. Гуссерля. Когнитивные образы, обнаруженные экспериментально в форме фреймов существуют в нашем сознании в неявном для нас виде и могут рассматриваться как форма невербализованного, но структурно организованного «неявного» знания (профессиональная интуиция, эмпирические навыки и т.д.). Теперь благодаря

информационным технологиям появилась возможность выявлять и управлять этими структурно организованными невербальными знаниями. Другими словами речь идет об использовании таких фундаментальных возможностей нашего сознания как категориальная интуиция, интенция, эйдетические и интенциональные объекты и т.д. Очень важно, что все перечисленные структуры и механизмы сознания уже получили инструмент для своего описания в соответствующих разделах математики, таких как теория топосов, алгебраическая топология, теория многообразий, объединяемых теорией категорий.

Ярким отражением происходящего прорыва является развитие когнитивной машинной графики. Открытие Б. Зенкиным и Б. Мандельбротом зрелищных геометрических феноменов, обладающих воздействием, как на эстетические, так и на эйдетические структуры нашего сознания и порождаемых вполне конструктивными явлениями окружающего мира позволило впервые однозначно связать строго рациональные знания с невербальными интенциональными знаниями, которыми мастерски «владеет» человеческое сознание. По сути, реализуется расширение и обновление поля рационального знания и этот прорыв способен революционно обновить весь арсенал технологий современной цивилизации.

Все выше высказанное можно проиллюстрировать на конкретной практически важной и актуальной проблеме визуализации терабайтных многомерных данных. Фактически, приведенное выше философски пафосное утверждение можно переформулировать следующим образом. Существует возможность расширения и обновление средств рационального анализа данных на основе генерации когнитивных образов.

6.2.1. Механизмы визуализации многомерных данных

Для решения задачи представления многомерных терабайтных данных в виде зрелищного образа в соответствии с изложенными выше идеями целесообразно напомнить, что при восприятии зрительных образов для сознания наиболее мощным является использование категориальной интуиции. Объекты, выявляемые на основе только категориальной интуиции, называются эйдетическими. Именно для этих объектов человеческое сознание способно осуществлять операцию выделения отдельных объектов, то есть их можно категоризировать. Эти объекты имеют фундаментальные свойства, которые улавливает человеческое сознание. Перечислим эти свойства, характеризующие отношения между объектами:

Рядоположенность и иерархическая упорядоченность объектов.

Отношение части и целого.

Принадлежность признака объекту.

Категориальные отношения.

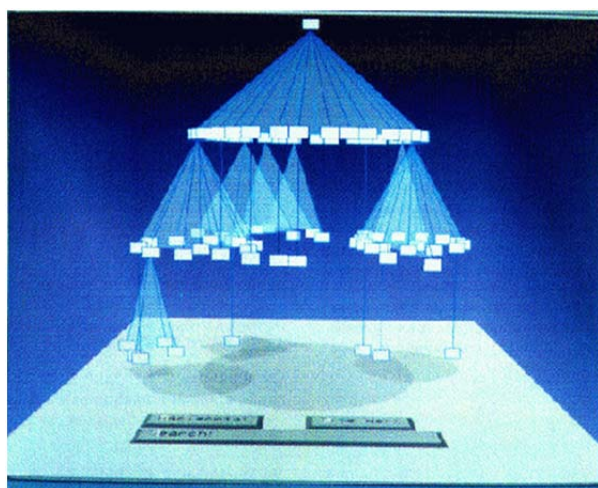
Возможность действий над объектами.

Выявление этих свойств нашим сознанием и есть проявление категориальной интуиции (усмотрение сущности). Поэтому принципиально важно [2–4] использовать структуры категориальной интуиции – рядоположенность, иерархичность объектов, отношение части и целого, что позволяет активировать интенциональность сознания для выявления скрытых категориальных отношений (в виде научных и инженерных метафор) между объектами.

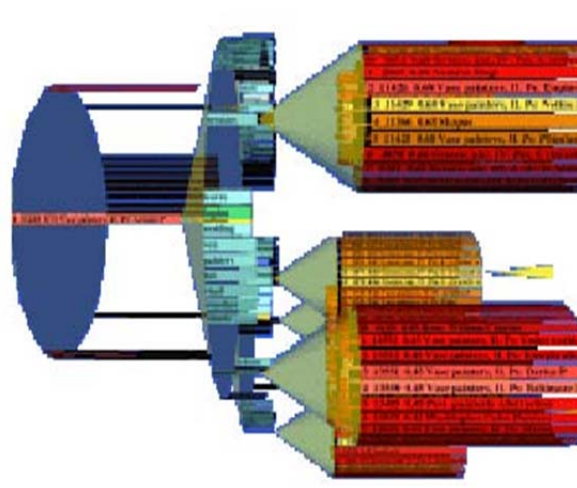
Прежде всего, для генерации когнитивных образов можно использовать свойство иерархической упорядоченности многомерных данных. Здесь можно отметить наиболее важные математические конструкции для множественных и перекрывающихся иерархических структур: иерархии, графы и сети (данные, содержащие узлы) и многомерные многообразия. Это помогает в поиске, классификации и манипулировании объектами, которые могут быть сгруппированы за счет скрытых алгебраических и топологических атрибутов. Существует два классических подхода для представления древовидных структур: подход связности узлов и контейнерный (вложенность). Эта метафора связности узлов в общем случае визуализирует узлы в пределах древовидной структуры, как небольшие метки и графически соединяет их связями, отражающими логические соединения между узлами. Метафора контейнера представляет дерево изображением потомка как меньшего контейнера в большем. Этот процесс выполняется далее рекурсивно, делением меньшего контейнера на контейнеры-потомки. Первоисточником здесь является парадигма информационного визуализатора Xerox PARC (Palo Alto Research Center). Древовидная структура с именем «Коническое дерево» (Cone Tree) представлена в трехмерном пространстве в попытке увеличить количество узлов, которые могут быть представлены на экране (рис. 1). Группы ссылок, исходящих из отдельного узла к его потомкам, образуют прозрачный конус с узлами потомков, равномерно распределенных по окружности основания конуса.

Категориальная интуиция инициируется благодаря возможности действий человека над когнитивными образами–метафорами. При выборе любого узла конусное дерево поворачивается так, что

выбранный узел выдвигается на передний план. Модель дисковых деревьев имеет два ключевых атрибута для связи с каждым конусом – точка ссылки и точка вершины. Точка ссылки – это узел над центральной точкой, лежащей на поверхности родительского узла. Точка вершины – это вершина конуса. При изменении расстояния между родительским узлом, точкой ссылки, точкой вершины и центральной точкой основания конуса можно получить различные когнитивные образы древовидных структур. При изменении этих расстояний конус может стать более плоским и превратиться в диск, что позволяет более эффективно использовать поверхность экрана. Рис. 2 показывает возможные для модели изменяемых дисковых деревьев конфигурации.



а) визуализатор с вертикальными осями вращения



б) визуализатор с горизонтальными осями вращения

Рис. 1. Визуализатор иерархической информации «Коническое дерево»

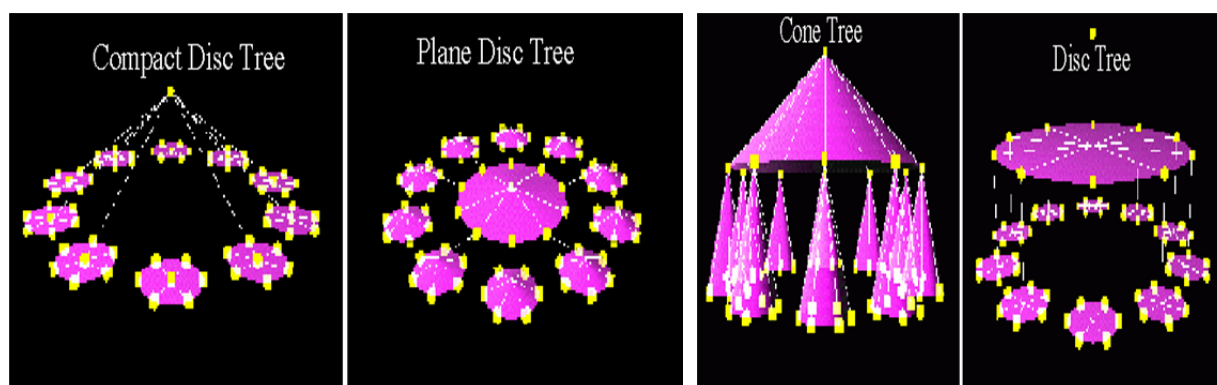


Рис. 2. Конфигурации для модели изменяемых дисковых деревьев

Гиперболический браузер. Гиперболический браузер использует технику наведения и контекста, основанную на гиперболической геометрии. Иерархия располагается равномерно на гиперболической плоскости и затем проецируется на поверхность круга для отображения на дисплее. Узлы в центре круга имеют наибольший размер, по мере приближения к краю круга их размер уменьшается, а количество растет экспоненциально (рис. 3).

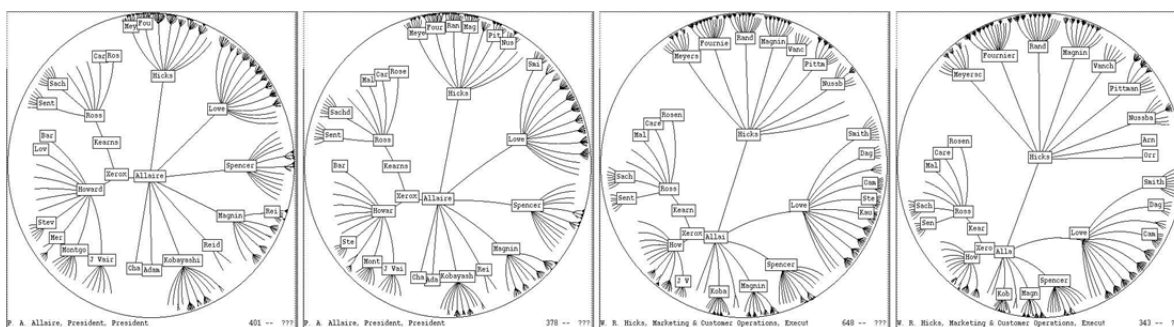


Рис. 3. Смена фокуса в гиперболическом браузере

Указанные два свойства – искажение типа «рыбий глаз» и возможность включения экспоненциально растущей структуры – главные аспекты данной метафоры. При выделении какого-либо подчиненного узла (смене фокуса), происходит плавная трансформация изображения, при этом выбранный узел располагается в центре круга, открывая обозримый вид для подчиненных ему узлов. В любом случае доступное пространство до какого-либо узла уменьшается, будучи непрерывной функцией его расстояния в дереве от центральной точки. Поскольку, контекст всегда включает некоторое количество родительских, дочерних узлов и узлов одного уровня, то это позволяет легко обследовать иерархию, не теряясь в ней. Гиперболический браузер поддерживает эффективный режим взаимодействия с иерархиями гораздо большего размера, чем обычные браузеры иерархий. Если обычный двухмерный браузер в окне 600x600 пикселей может отобразить 100 узлов, то гиперболический браузер может отобразить 1000 узлов, из них около 50 находятся в фокусе и легко читаемы.

«Информационный куб». Метафора иерархической вложенности в трехмерном пространстве была использована в модели «Информационный куб». Этот подход использует полупрозрачные кубы, вложенные один в другой по принципу матрешки, чтобы выразить эффект ветвления узлов (рис. 4). Уровень прозрачности кубов может регули-

роваться, в зависимости от того, насколько глубоко необходимо просматривать иерархическую структуру той или иной ветви. При этом другие ветви (кубы) можно прятать (делать прозрачными), чтобы избежать визуального хаоса.



Рис. 4. Модель «Информационный куб»

«Динамическое проецирование» из многомерных многообразий, построенных на многомерных данных и дополненное масштабированием с использованием инвариантных робастных статистик дополняет приведенный список когнитивных инструментов [4]. На рис. 5 представлен когнитивный образ, построенный на основе терабайтного многомерного каталога галактик HUDF.

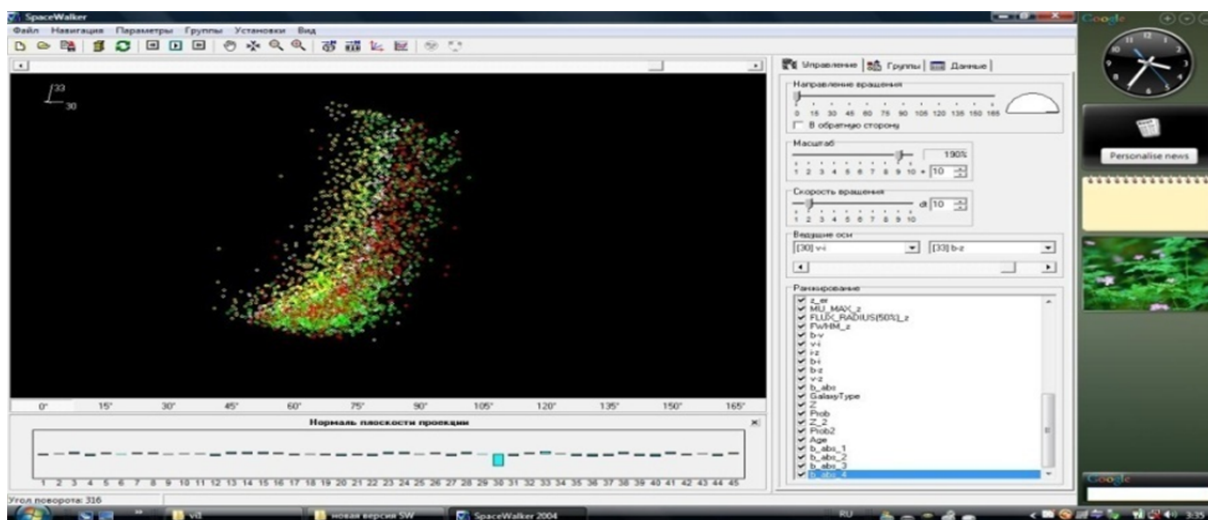


Рис. 5. Когнитивный образ совокупности объектов (каталог галактик HUDF) который состоит из точек¹ в 44-мерном пространстве

¹ Каждая из этих точек представляет отдельную галактику.

6.2.2. Пример когнитивной визуализации многомерных терабайтных данных астрофизического обзора COSMOS

Одной из важных задач анализа данных космологических обзоров является выявление крупномасштабных неоднородностей в распределении галактик [7]. В данной работе на основе использования непараметрического теста Вилкоксона и когнитивной визуализации статистических многомерных контрастов [4] с учетом измерения пространственной корреляционной функции галактик проведено сравнение и выявление статистически значимых контрастов наблюдаемых пространственных неоднородностей распределения галактик на масштабах сотен мегапарсек. Для этого были использованы наблюдаемые отклонения от однородного распределения приведенные в работе [7]. На фиксированных пространственных участках, в которых наблюдениями заданы выборки отклонений, тестами Вилкоксона и Ходжеса Лемана проверялись и были выявлены статистически значимые контрасты. Вычисленный достигаемый уровень значимости (P-VALUE), характеризующей достаточно надежное выявление эффекта, имеет величину 0,043. Отмеченные зоны неоднородности выявляются и при когнитивном анализе всей многомерной совокупности пространственных участков при разных масштабах (рис. 6).

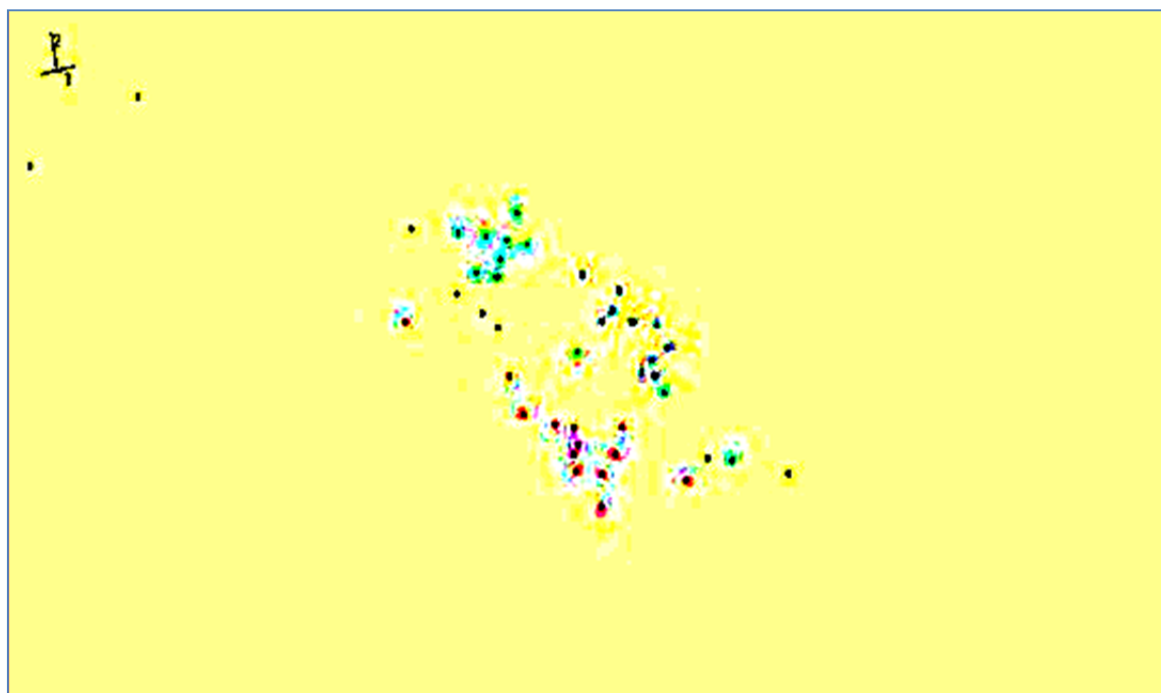


Рис. 6. Когнитивный образ интегрированных подсчетов на основе обзора COSMOS

Таким образом, приведенные выше механизмы генерации когнитивных образов отличаясь математическими подходами (которые включают методы топологии, современной геометрии и алгебры) опираются на идеи гносеологии, когнитивной психологии и технологии искусственного интеллекта и показывают практическую эффективность и полезность для визуализации терабайтных данных. Из иностранных корпораций, наибольший успех достигнут в подобных разработках в компании Facebook (См. рис. 7). Перспективное сотрудничество в области когнитивных технологий, организует лаборатория компьютерных наук и искусственного интеллекта (CSAIL) Массачусетского технологического института [1] .



Рис. 7. Когнитивный образ работы сайтов в интернете

6.3. Пример выявления аномальных ситуаций при мониторинге сложных систем на основе когнитивных технологий

Процесс мониторинга аномальных ситуаций в сложных системах обеспечивает своевременное обнаружение катастрофических или опасных ситуаций в поведении этих систем [1]. Подобного рода мониторинг обеспечивает достаточно устойчивое и надежное управление сложными системами [1–4].

Одним из успешных инструментов такого мониторинга является когнитивная визуализация поведения сложной системы [2, 3]. Рассмотрим пример использования системы когнитивной визуализации SW для изучения поведения сложной системы движения подвижного состава РЖД.

В связи с возрастающими требованиями и объемами перевозок необходим своевременный и качественный контроль за работой важнейших

сортировочных и грузовых станций. Рациональное решение задач по работе грузовых станций должно обеспечить минимальные расходы, связанные с простоем вагонов, маневровых локомотивов, погрузочно-разгрузочных машин, а также с их эффективной эксплуатацией в перевозочном процессе. Для контроля за выполнением установленных показателей и для оценки работы станции в РЖД установлены системы и формы учета работы станций, которые осуществляют формирование важнейших показателей работы и представление отчетности (ДУ, ДО, ГУ, ГО).

С целью разработки мер по улучшению работы станции периодически выполняется когнитивный анализ [3]. Он позволяет вскрыть резервы и найти «узкие» места в работе отдельных участков станции, при этом выявляется ряд негативных факторов, такие как, большие простои, которые ведут к снижению перерабатывающей способности. Порядок исследования когнитивного образа:

1. Выявление характерных станций в норме (пары параметров).
2. Выявление отклонений от нормы.
3. Выявление аномальных объектов по всем n-параметрам (вращение образа).
4. Выявление динамики.
5. Выявление зависимостей по парам данных.

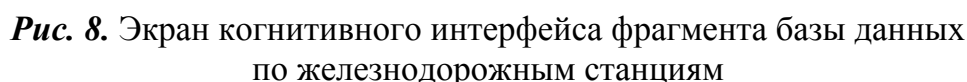
6.3.1. Описание методики выявления аномалий

Выявление наиболее производительных станций (для пары параметров). Выбираем параметры «отправлено-прибыло». Выбираем наиболее производительные станции (красим в красный цвет) и смотрим их в таблице данных, проверяя правильность визуальной сегментации (рис. 8).

Выявление станций среднего уровня производительности (пары параметров) для той же пары «отправлено-прибыло». Отмечаем норму одним цветом и смотрим их в таблице данных, проверяя правильность визуальной сегментации.

Выявление отклонений от нормы. Для той же пары «отправлено-прибыло». Здесь две сегментации. Белым цветом: «много прибыло мало отправлено». Смотрим их в таблице данных, проверяя правильность визуальной сегментации и видим, что в эту группу попали припортовые станции. Вторая сегментация – это малопроизводительные станции – другой цвет.

Выявление аномальных объектов по всем n-параметрам (за счет когнитивного вращения многомерного образа – когнитивный режим).



256

Аналитическая справка о работе грузовых станций (ОЭ)

Год: 2011
Месяц: Октябрь

Дорога	Станция	Прибыло поездов	Отправлено поездов		Вагонооборот	Рабочий парк вагонов	Погружено вагонов	Выгружено вагонов	Простой местного вагона									Простой местного вагона на ответственности ОАО "РЖД"				
			Всего	В т.ч. своего формирования					Всего	В том числе							Всего	В том числе				
										Вагонов резерва	Приватных вагонов на путях отстоя	Прочий простой местных вагонов	Вагонов, участвующих в грузовых операциях	В том числе				от прибытия до подачи	грузовые операции	ожидающие уборки грузовых фронтов	ожидающие отправления	
														от прибытия до подачи	грузовые операции	ожидающие уборки грузовых фронтов						
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1 АПАТИТЫ	ОКТ	15	17	12	1728	1839	205	18	120,04	0,00	0,00	6,72	113,32	64,47	23,83	10,93	14,09	89,52	66,01	0,49	9,07	13,95
2 ОЛЕНЕГОРСК	ОКТ	23	25	6	2770	825	225	19	45,77	0,00	0,00	1,59	44,18	13,20	18,11	4,84	8,03	24,23	13,76	0,18	2,05	8,24
3 МУРМАНСК	ОКТ	13	12	12	1413	1144	23	651	37,59	0,00	0,00	1,00	36,59	12,20	5,90	12,33	6,16	20,93	12,16	0,15	2,93	5,69
4 ВЫСОЦК	ОКТ	7	7	7	814	749	0	397	41,58	0,00	0,00	0,00	41,58	3,64	20,37	6,97	10,60	14,75	3,64	0,00	0,84	10,27
5 КАМЕННОГОРСК	ОКТ	6	6	6	626	475	239	0	42,27	0,00	0,00	1,05	41,22	8,93	20,03	3,95	8,31	21,17	8,95	0,08	3,82	8,32
6 КУЗНЕЧНОЕ	ОКТ	10	11	4	1021	506	188	0	50,52	0,00	0,00	1,92	48,60	3,44	25,26	0,42	19,48	24,27	4,51	0,00	0,28	19,48
7 КОСТОМУК-ТОВ	ОКТ	8	8	8	899	721	411	29	38,28	0,00	0,00	2,33	35,95	9,28	15,52	7,78	3,37	20,22	10,68	0,27	5,79	3,48
8 АВТОВО	ОКТ	16	17	17	1737	1243	84	810	36,24	0,00	0,00	0,43	35,81	6,56	15,68	7,40	6,17	13,98	6,64	0,05	1,18	6,11
9 НОВЫЙ ПОРТ	ОКТ	9	11	11	926	1262	98	427	63,43	0,00	0,00	0,45	62,98	9,53	21,10	18,14	14,21	26,96	9,71	0,31	2,77	14,17
10 НОВГОРОД-НА-	ОКТ	6	6	4	625	1475	152	67	210,52	0,00	0,00	17,80	192,72	11,93	149,59	24,07	7,13	21,80	12,05	0,53	2,05	7,17
11 КИРИШИ	ОКТ	15	14	14	1710	1572	728	80	42,21	0,00	0,00	2,82	39,39	11,39	14,01	9,10	4,89	22,52	11,49	0,31	5,74	4,98
12 ЛУЖСКАЯ	ОКТ	18	18	18	2241	1478	6	1041	28,58	0,00	0,00	0,13	28,45	8,99	7,52	3,91	8,03	13,87	8,98	0,00	0,10	4,79
13 КАЛИНИНГ-СОР	КЛГ	19	22	21	1694	1028	18	176	55,83	0,00	0,00	1,51	54,32	10,51	13,48	17,51	12,82	28,08	10,50	2,06	3,24	12,28
14 БАЛТИЙСК ЛЕС	КЛГ	2	2	2	181	343	1	87	82,74	0,00	0,00	0,19	82,55	2,41	48,15	20,54	11,45	20,50	2,60	7,49	0,00	10,41
15 АЗОТНАЯ	МОСК	2	3	3	191	623	68	16	186,71	0,00	0,00	9,99	176,72	22,22	121,97	29,40	3,13	46,29	22,91	0,00	17,15	6,23
16 МАЛЬЧИКИ	МОСК	4	4	3	192	225	8	35	134,18	0,00	0,00	1,66	132,52	16,37	47,96	57,40	10,79	52,71	18,04	4,99	18,91	10,77
17 ЯНИЧКИНО	МОСК	6	6	6	672	887	229	110	60,89	0,00	0,00	0,67	60,22	15,68	27,37	11,67	5,50	23,99	15,75	0,32	2,77	5,15
18 МИХАЙЛОВ РУД	МОСК	6	12	12	714	653	156	14	47,54	0,00	0,00	10,63	36,91	8,10	8,20	16,33	4,28	23,67	8,59	0,08	10,54	4,46
19 КУРБАКИНСКАЯ	МОСК	9	18	18	1401	1076	548	15	36,73	0,00	0,00	2,47	34,26	7,36	13,82	6,74	6,34	18,49	8,06	0,18	3,80	6,45
20 ПРИСАДЫ	МОСК	6	6	6	570	1005	124	215	97,68	0,00	0,00	1,80	95,88	9,40	48,48	20,79	17,21	27,27	9,09	0,11	1,88	16,19
21 ЛЕСОК	МОСК	2	2	2	219	366	13	52	102,03	0,00	0,00	1,01	101,02	13,47	38,10	32,41	17,04	47,77	13,54	10,03	7,24	16,96
22 СТЕЬКИНО II	МОСК	10	11	11	1235	1620	529	39	57,54	0,00	0,00	8,08	49,46	25,10	15,22	8,55	0,59	36,63	28,81	0,79	6,43	0,60
23 СЕВЕРНАЯ	МОСК	3	3	3	279	824	105	12	119,99	0,00	0,00	8,66	111,33	7,92	90,91	9,33	3,17	18,97	7,58	0,03	7,51	3,85
24 ВОСКРЕСЕНСК	МОСК	10	11	7	898	1357	59	107	157,68	0,00	0,00	16,33	141,35	15,48	54,90	64,51	6,46	41,76	18,32	1,44	12,54	9,46
25 НАВАШИНО	ГОРЬК	4	5	4	491	2206	103	143	226,25	0,00	0,00	15,46	210,79	23,68	125,58	49,99	11,54	37,11	23,56	0,81	1,50	11,24
26 НУЯ	ГОРЬК	5	5	3	489	541	125	13	80,78	0,00	0,30	0,76	79,72	5,11	3,35	60,36	10,90	17,15	5,18	1,00	0,06	10,91
27 ПРАВДИНСК	ГОРЬК	3	4	4	204	273	18	71	76,84	0,00	0,00	0,81	76,03	4,66	36,27	25,90	9,20	16,12	4,62	0,05	1,67	9,78
28 БЕЛЕЦИНО	ГОРЬК	11	12	12	1295	2502	560	78	87,26	0,00	0,00	1,69	85,57	41,20	17,12	15,85	11,40	51,49	37,40	1,81	1,77	10,51
29 ЧЕПЕЦКАЯ	ГОРЬК	3	4	4	367	1557	135	40	188,22	0,00	0,00	12,38	175,84	5,33	21,14	133,49	15,88	22,39	5,36	2,13	0,09	14,81
30 НИЗОВКА	СЕВ	4	4	3	408	950	66	108	113,60	0,00	0,00	2,22	111,38	0,37	82,50	25,42	3,09	8,69	0,87	0,00	4,17	3,65
31 ИНТА II	СЕВ	5	4	4	202	290	86	30	60,84	0,00	0,00	0,68	60,16	6,97	21,06	4,44	27,69	37,84	6,97	0,62	2,15	28,10
32 МУЛЬДА	СЕВ	5	4	4	441	601	179	3	45,78	0,00	0,03	3,05	42,70	5,65	24,56	2,30	10,19	17,72	5,48	0,00	2,04	10,20
33 ИСАКОГОРКА	СЕВ	19	19	16	1741	1190	45	190	69,89	0,00	0,00	2,25	67,64	7,56	19,61	29,83	10,64	21,39	7,52	0,01	2,80	11,06
34 ЧЕРЕПОВЕЦ II	СЕВ	34	27	11	3965	5787	394	848	124,40	0,00	0,00	0,81	123,59	10,29	99,06	11,98	2,26	18,56	10,32	1,04	4,94	2,26
35 КОШТА	СЕВ	41	47	19	5915	3118	263	329	89,63	0,00	0,00	6,47	83,16	8,55	49,92	14,75	9,94	23,94	8,25	0,19	1,33	14,17
36 СИЛЬНИЦЫ	СЕВ	2	3	2	220	190	67	1	63,67	0,00	0,00	1,59	62,08	15,03	20,68	9,47	16,90	41,50	16,37	0,01	8,11	17,01
37 НОВОЯРОСЛАВС	СЕВ	10	10	10	1241	1316	574	21	49,11	0,00	0,00	1,72	47,39	24,59	12,45	4,92	5,43	35,05	25,24	0,07	4,09	5,65
38 ТАГАНРОГ	С-КАВ	28	28	11	3104	1888	67	205	124,56	0,00	0,00	6,25	118,31	24,25	28,24	40,17	25,65	56,88	27,31	4,61	5,00	19,96
39 БЕЙСК	С-КАВ	2	1	1	76	136	9	31	76,73	0,00	0,00	3,89	72,84	19,72	24,94	12,63	15,55	42,61	18,79	4,94	3,61	15,27
40 НОВОРОССИЙСК	С-КАВ	24	21	21	2683	3083	183	1241	55,64	0,00	0,00	0,22	55,42	24,14	9,69	8,65	12,94	41,18	24,09	0,82	4,13	12,14
41 ГРУШЕВАЯ	С-КАВ	6	6	6	639	283	85	317	22,55	0,00	0,00	0,00	22,55	9,23	2,68	4,34	6,30	16,92	9,23	0,30	1,27	6,12

Рис. 9. Таблица многомерных данных

Таким образом, когда мы видим выпадающие объекты-точки, то это означает, что выделив (раскрасив эти точки) мы можем вернуться в таблицу данных и найти тип аномальных параметров, которые привели к выбросу этой точки.

Таким образом, когнитивный образ позволяет мгновенно выделить аномальный параметр (тип характеристики для данной станции), который возможно влияет на всю ситуацию. Следует особо отметить, что быстро выявить эту аномальную характеристику и, соответственно, аномальную станцию, которая влияет на весь многомерный набор данных, просто анализируя многомерную таблицу данных, практически невозможно. Пример показан на рис. 9.

Визуальное наблюдение псевдо-трехмерного когнитивного образа показывает явно выпадающие из облака точки, которые фиксируются желтым цветом. В таблице данных по цвету находим две вполне рядовых станции и, просматривая всю строчку параметров, именно для этих станций обнаруживаем, что они обладают аномальными значениями по ряду параметров простоя вагонов и показателями прибыло-убыло. Заметим, что такая когнитивная селекция быстро работает при десятках тысяч объектов и сотнях параметров этих объектов.

6.3.2. Выявление динамики

Если наборы многомерных данных изменяются (например, раз в день), то генерация когнитивных образов, описанная выше, для сменяющихся наборов данных порождает когнитивный образ, который день ото дня изменяется. И в случае серьезных изменений в работе множества станций когнитивный образ может радикально измениться! Эти изменения могут быть принципиально радикальными с визуальной точки зрения, а значит, появляется возможность человеку оперативно обнаружить аномальные изменения в работе этих станций. Это уже является сильным подспорьем для визуального быстрого контроля за всей совокупностью (десятки тысяч) станций – объектов. Пример показан на рис. 10. Здесь показан когнитивный образ, порожденный описанным выше набором данных. Но теперь этот набор данных предъявляется для нескольких дней. Наблюдая когнитивный образ для разных дней можно наглядно видеть, как он меняется!!

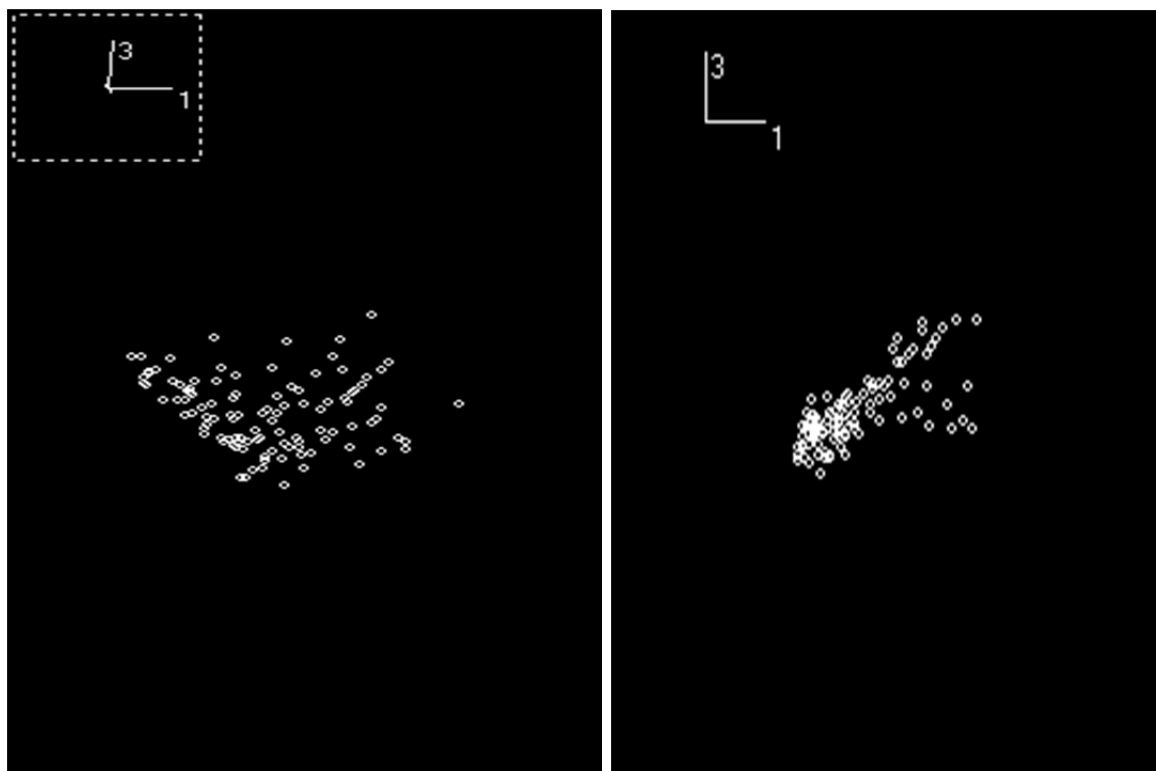


Рис. 10. Ракурсы когнитивного образа многомерных данных

6.3.3. Выявление зависимостей по парам данных

Еще очень важный ресурс генерации когнитивных образов многомерных данных состоит в том, что пользователь получает возможность количественно и визуально описывать как двумерные, так и многомерные функции зависимости параметров станций между собой! Эти функции зависимости могут быть, как линейными, так и нелинейными. Например, для линейных двумерных зависимостей достаточно, используя программу генерации когнитивных образов, подобрать для интересующих пользователя пар параметров вид когнитивных образов в виде прямой, наклоненной под углом относительно осей этих параметров. Например, рис. 11. Здесь представлены зависимости параметров: простой местных вагонов (9) и простой вагонов, участвующих в грузовых операциях (12). Сразу видно, что зависимость имеет линейный характер. Данную зависимость можно сразу использовать для вычисления зависимостей между этими параметрами. Тангенс угла наклона говорит о силе связи между параметрами. Если пользователь интересуется многомерными особенностями этой зависимости, достаточно включить движение облака в многомерном пространстве.

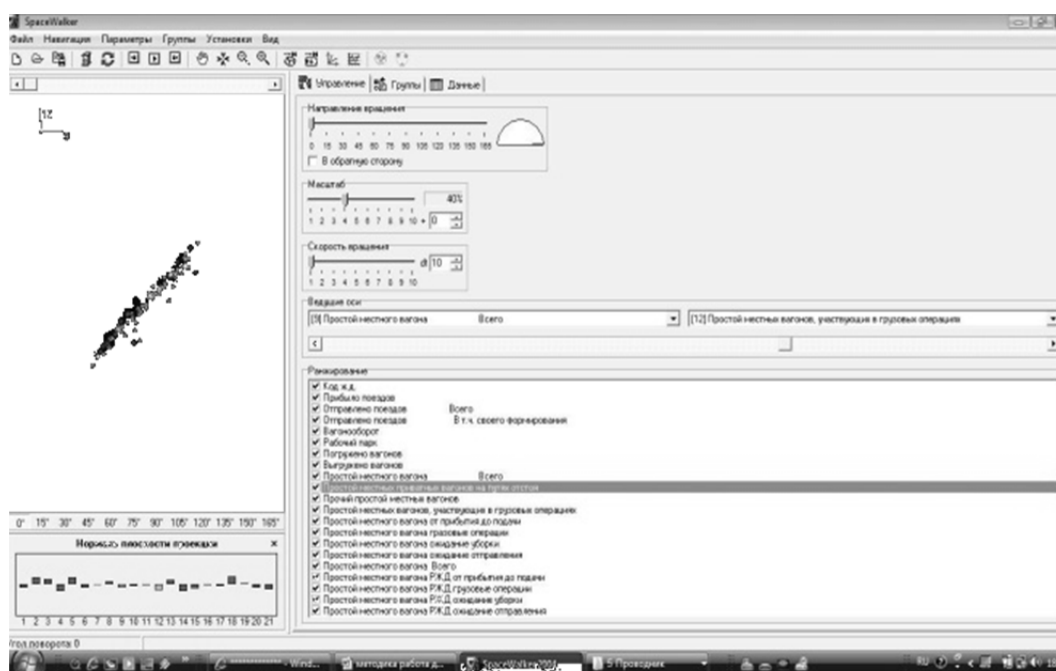


Рис. 11. Когнитивный образ принимает вид узкого эллипсоида, что «сигнализирует» о регрессионной зависимости

Пользователь увидит псевдо-трехмерное изображение поверхности этой зависимости. Видно, что эта зависимость носит характер многомерной плоскости (т.е. остается линейной). Пример нелинейной зависимости представлен на рис. 12.

После наблюдения и использования полученных зависимостей можно обратиться к возможностям современных программ многомерного анализа, которые позволяют получить соответствующие аналитические и математические описания найденных данных.

Подчеркнем особенности применения пакетов стандартного многомерного анализа. Прежде всего, перед использованием конкретных программ стандартного многомерного анализа требуется заранее, до применения этих программ, задать модель данных. Линейная модель, нелинейная, тип распределения данных и т.д. Если нелинейная, то нужно задать уравнение нелинейной зависимости, которая и будет использоваться при вычислении. Если задана модель неверно, все вычисления недостоверны.

Применение описываемых программ когнитивной визуализации не требует таких модельных предположений. Наоборот, применение описываемых программ когнитивной визуализации позволяет сначала увидеть вид зависимости, а потом осознано перейти к стандартному

многомерному статистическому анализу, тем самым повышая достоверность его работы.

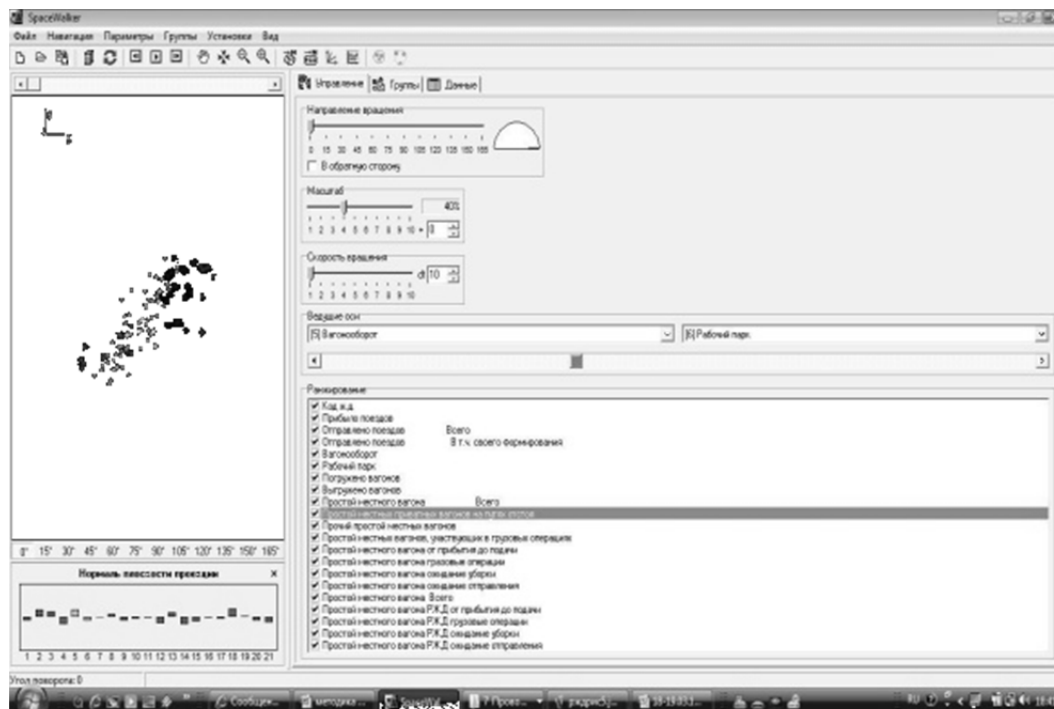


Рис. 12. Когнитивный образ принимает вид псевдо-трехмерного «полотенца», что «сигнализирует» о нелинейной регрессионной зависимости

6.3.4. Сравнение с другими методами

Можно сравнить предлагаемые инструменты когнитивной визуализации с достаточно традиционными графическими средствами (например столбчатыми диаграммами). Их использование дает возможность фиксировать частотные свойства только по двум параметрам. Подобные диаграммы не фиксируют даже трехмерные зависимости.

Заключение по п. 6.3

Предлагаемые когнитивные инструменты и методики наглядно и просто фиксируют не только частотные двумерные закономерности, но и статистические многомерные зависимости.

Кроме того, предлагаемые когнитивные инструменты позволяют увидеть и оценить многомерные регрессионные, корреляционные и частотные свойства многомерных данных.

Кроме того, эти инструменты обеспечивают оперативное наблюдение и обнаружение аномальных и стандартных ситуаций в много-

мерных данных. Они позволяют проводить визуальную классификацию многомерных данных и выявление главных факторов

6.4. Технология и методика генерации псевдо 6-D образов

Разработана принципиально новая технология и методика когнитивно-визуального анализа больших объёмов гетерогенных и отягощённых ошибками данных основанная на использовании когнитивных свойств человеческого сознания [1–9]. Модернизирована и инсталлирована система SpaceWalker на высокопроизводительных графических станциях. Тестирование разработанной ранее программной системы когнитивной визуализации SpaceWalker показало перспективность развития метода для анализа данных различной природы и форматов.

С использованием опыта разработки программной системы SpaceWalker разработана система SpaceHedgehog, позволяющая непосредственную динамическую визуализацию объектов размерности более 3D. Примеры интерфейсов показаны на рис. 13–15. На основе исследования и развития технологии когнитивной визуализации разработана методика и технология когнитивной 6-D визуализации многомерных данных. Здесь используются мнемонические возможности когнитивных ресурсов человеческого сознания. Выполнена разработка мультиплатформенных Java-версий систем SpaceWalker и SpaceHedgehog.

Применение методов 6-D когнитивной визуализации существенно расширяет возможность оценки, отбора и объединения данных для анализа больших коллекций (архивов) разнородных данных и данных, отягощённых ошибками.

Использование методов 6-D когнитивной визуализации помогает в решении проблем совместной обработки и анализа разнородных данных, равно как данных с ошибками, пропусками и искажениями, является особенно актуальной для фундаментальной экспериментальной науки и, в частности, астрофизики.

Состояние современной прикладной информатики определяется во многом уровнем качества машинной графики для систем поддержки интеллектуальных решений и средств мониторинга и контроля.

Прежде всего, машинная графика должна обеспечивать повышение таких важнейших показателей прикладных информационных систем как надёжность и достоверность результатов работы решающих процедур контроля и достоверность, поступающих данных.

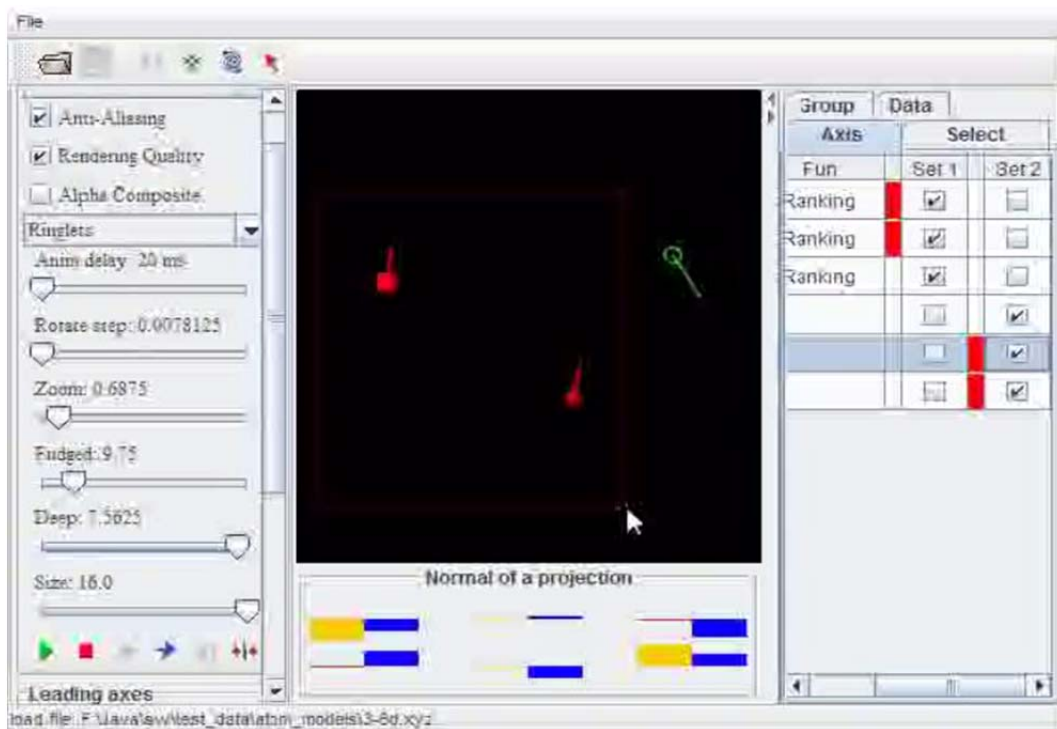


Рис. 13. Модельные многомерные данные визуализируются как псевдо 6-D объекты

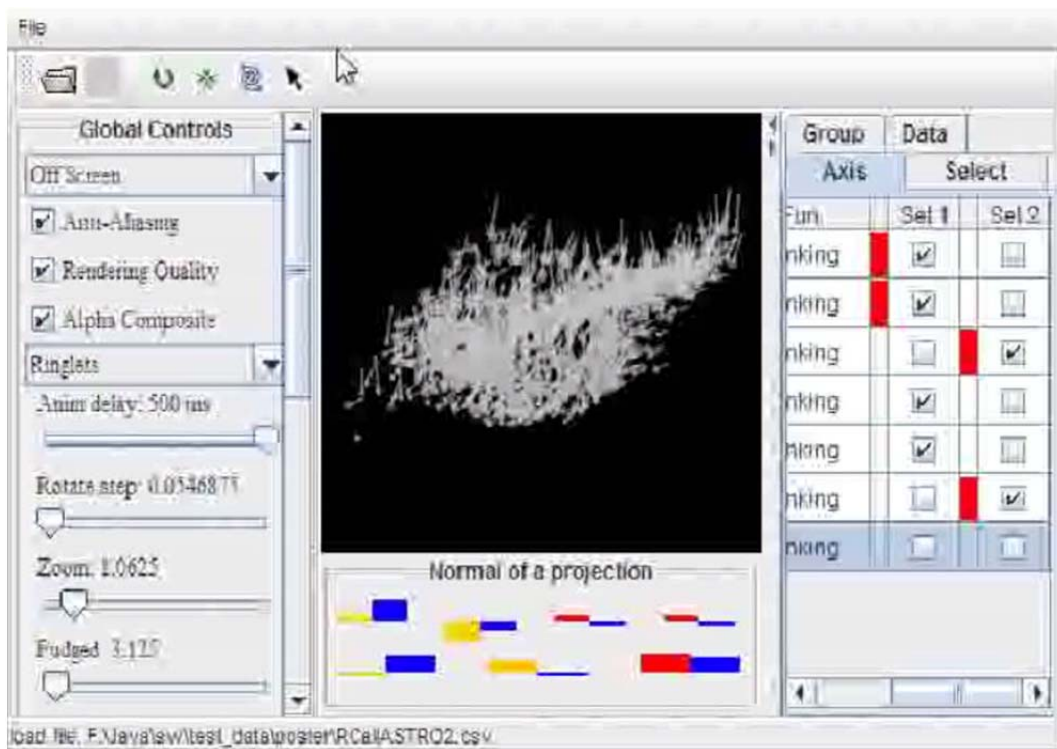


Рис. 14. Реальные многомерные данные радиоастрономического каталога визуализируются как псевдо 6-D объекты

Предложенный метод 6-D визуализации позволяет осуществлять подбор наилучшего положения когнитивного образа самому пользователю. При этом пользователь, опирается на свою интуицию и когнитивный образ перед глазами. Имея возможность активно влиять на ориентацию плоскости проекции в многомерном пространстве, исследователь свободен от предварительных соображений о статистической (геометрической) структуре данных, которые представляют объекты. Человек непосредственно видит на экране проекции кластеров или многомерных поверхностей, в которые формируются его данные. И этот зрелищный образ стимулирует его интуитивное понимание исследуемых объектов и формирование научных метафор. Особенно это важно при анализе статистических связей, факторном анализе и при распознавании образов.

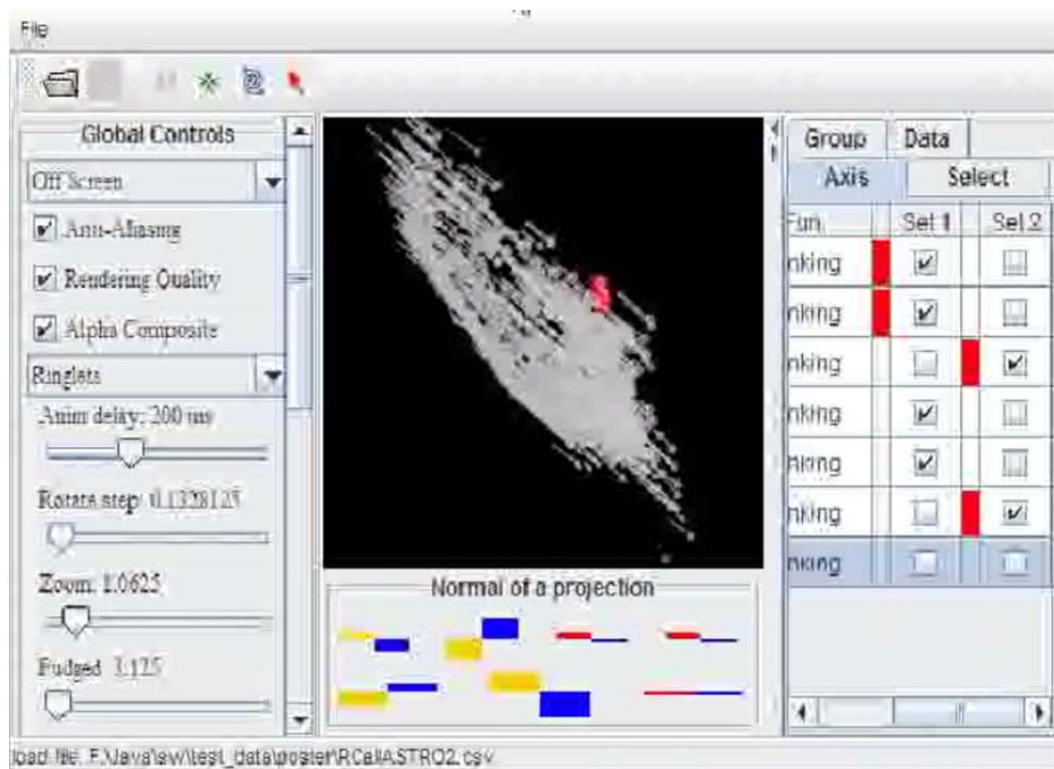


Рис. 15. Результаты выявления особых объектов в многомерных данных радиоастрономического каталога после 6-D визуализации

Тот факт, что процедура динамической визуализации не опирается на априорные сведения о природе объектов, а значит и не привносит в проекции искажающих влияний той или иной модели, дает возможность использовать визуализированные образы в условиях глубокой априорной неопределенности.

6.5. Выявление экстремальных состояний сложных природных и технических систем на основе когнитивного мониторинга многомерных временных рядов

В настоящее время регулярно возрастает количество чрезвычайных ситуаций на природных и промышленных объектах, происходит повреждение дорогостоящего оборудования, загрязнение окружающей среды и т.д. [1]. Для своевременного выявления и предупреждения чрезвычайных ситуаций необходимо обеспечить обнаружение аномальных состояний объектов. Одной из сложных предметных областей, с точки зрения своевременного выявления возможности возникновения чрезвычайных ситуаций, является энергетика. Мониторинг состояния нефтегазодобывающих платформ осуществляется на основе данных, поступающих с различных встроенных устройств, а также специализированных внешних датчиков и систем контроля состояния объектов. Состав регистрируемых многомерных данных позволяет контролировать состояние механических элементов, параметры вибраций и динамических воздействий, а также ряд других показателей, характеризующих техническое состояние анализируемых объектов. Например, это набор многомерных характеристик силовых агрегатов ГРЭС. Так для Красноярской ГРЭС осуществляется мониторинг более двадцати параметров для всех гидроагрегатов. В процессе мониторинга регистрируются изменения этих характеристик во времени в виде многомерных временных рядов. Однако существуют серьезные трудности решения задач выявления чрезвычайных ситуаций на сложных технических объектах. Прежде всего, это тот факт, что наблюдаемые многомерные ряды обладают плохо прогнозируемым поведением. Еще одна трудность заключается в многомерном характере временных рядов, которые в сочетании с большим числом характеристик объектов и объемов рядов затрудняют измерение взаимных корреляций и ковариаций. Преодоление этих трудностей возможно средствами когнитивной компьютерной графики [2]. Когнитивная компьютерная графика позволяет представить многомерные временные ряды в виде когнитивных псевдотрехмерных образов, которые стимулируют эмпирическую интуицию человека-оператора и появляется возможность обнаруживать синхронные взаимные корреляции и тренды в многомерных данных, с последующей их объективацией.

В данной работе приводятся модернизированные алгоритмы динамического проецирования многомерных временных рядов, обеспе-

чивающие генерацию когнитивных образов многомерных рядов. Когнитивные образы позволяют визуализировать многомерные взаимные корреляции и тем самым обнаруживать экстремальные состояния объектов мониторинга. Хочется отметить, что практика применения когнитивной машинной графики показала, что малые, но синхронные изменения характеристик носят синергетический характер и часто являются предвестниками чрезвычайных ситуаций.

6.5.1. Алгоритмы когнитивной визуализации многомерных временных рядов

Обозначим набор из p временных рядов в виде d_j ($j=1, \dots, p$), наблюдаемого в n временных точках. Значение d_j во временной точке t обозначается $\{d_{ij}\}$, а совокупность (p, n) значений d матрицей D . Как и одномерном случае D рассматривается как реализация случайного процесса. В общем случае, вектор строка из D может содержать тренд, сезонность и осцилляции. Многомерные ковариации и корреляции, а также взаимные многомерные ковариации и корреляции в виде квадратных матриц описывают зависимости внутри рядов и их синхронизмы для любых запаздываний s . Пусть входные данные описываются матрицей:

$$D = \{d_{ij}\} ; i = (1, n), j = (1, p),$$

где d_{ij} – действительные значения переменных (измеряемых параметров, признаков); n – число актов измерения всех p параметров объекта в последовательности времени наблюдения; p – число измеряемых параметров (признаков) объектов наблюдения. Многомерный временной ряд представляется в виде облака n точек в p -мерном пространстве параметров. Далее, будем полагать это p -мерное пространство параметров евклидовым RP , что упрощает дальнейшие рассуждения (рассмотренные далее алгоритмы остаются рабочими в подходящих базисах аффинного пространства и ряда других базисов).

Построим отображения Φ как проекцию этого облака на двумерную плоскость $Q2$, проходящую через начало координат пространства RP . Пусть в $Q2$ заданы единичные ортогональные вектора u и v , используя их несложно вычислить координаты (x, y) проекции данных на двумерную гиперплоскость Q :

$$x_i = pr_u \mathbf{d}_i = \mathbf{d}_i \cdot \mathbf{u}, \quad y_i = pr_v \mathbf{d}_i = \mathbf{d}_i \cdot \mathbf{v}, \quad (1)$$

где исходная матрица данных D описывает облако точек T в многомерном пространстве RP .

Предлагается алгоритм, который строит непрерывную последовательность положений $Q2$, образующих «траектории», вдоль которых и отслеживается динамика образа. Для этого строится процедура динамического вычисления последовательности пар векторов $\{\mathbf{u}, \mathbf{v}\}$. Каждая ортогональная пара векторов $\{\mathbf{u}, \mathbf{v}\}|g$ будет определять двумерную плоскость $Q|g$ и ее ортогональный базис, где g набор управляющих параметров небольшой размерности. Определим гиперплоскость W (размерности $p - 1$), проходящую через начало координат пространства RP нормальным уравнением:

$$\mathbf{x} \cdot \mathbf{n} = 0, \quad (2)$$

где $\mathbf{n} = \{n_j\}$ – вектор нормали, $\mathbf{x} = \{x_j\}$ – независимые переменные p -мерного пространства. Используя приведенные формулы осуществляется динамическая проекция данных D на двумерную плоскость. Например, выбрав ведущие оси, задав нормаль и направление вращения, изменяя с небольшим приращением угол поворота φ , получим динамический циклический обзор данных.

6.5.2. Практическое применение когнитивных технологий для выявления аномальных состояний объектов мониторинга

Для предварительного когнитивного анализа выделен ряд параметров гидроагрегата Красноярской ГРЭС (например: расход, мощность реактивная, мощность активная, вибрация турбинного подшипника, вибрация ого подшипника и т.д. 22 параметра). После первичного рассмотрения оператором когнитивного образа объединённого 22-х мерного массива данных, были выбракованы и исключены из дальнейшего рассмотрения параметры 7 и 11, не меняющие значение в указанный период времени. Дальнейший анализ показал наличие в рассматриваемой предметной области трех аномальных статистических зависимостей и особенностей способных влиять на поведение гидроагрегатов системы в целом. На приведённых ниже скриншотах выделена одна такая особенность, способная оказать влияние на параметры 1 и

2, вызванная возможно частичным внутренним нарушением обычного функционирования системы или возможным внешним, по отношению к системе, воздействием (рис 16).

Выявлены временные зоны для резкого изменения ряда параметров гидроагрегата. Эти параметры существенным образом определяют состояние агрегата (рис 17).

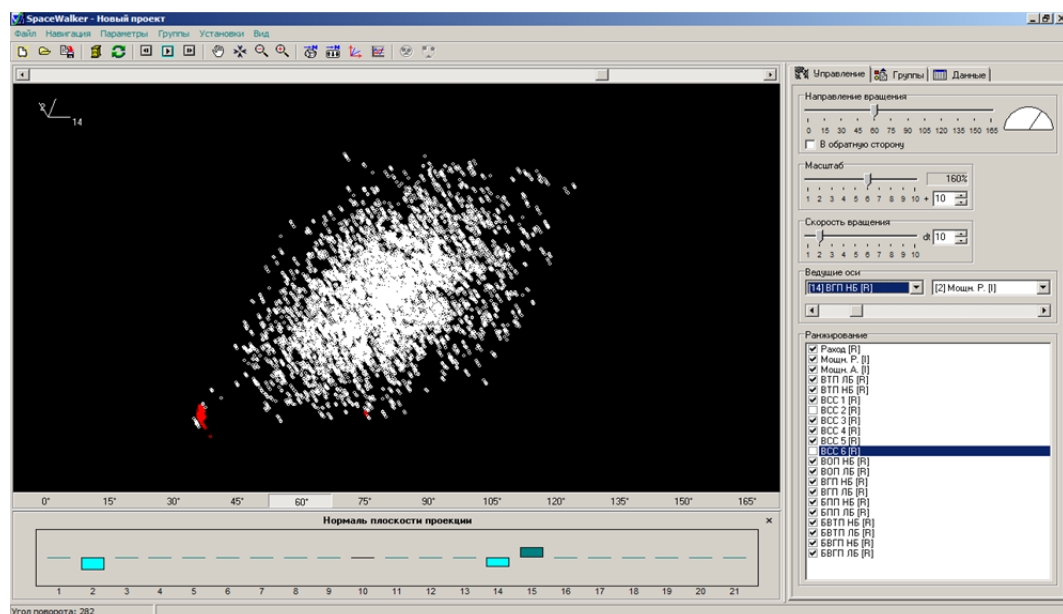


Рис. 16. Временные зоны резкого изменения ряда параметров гидроагрегата

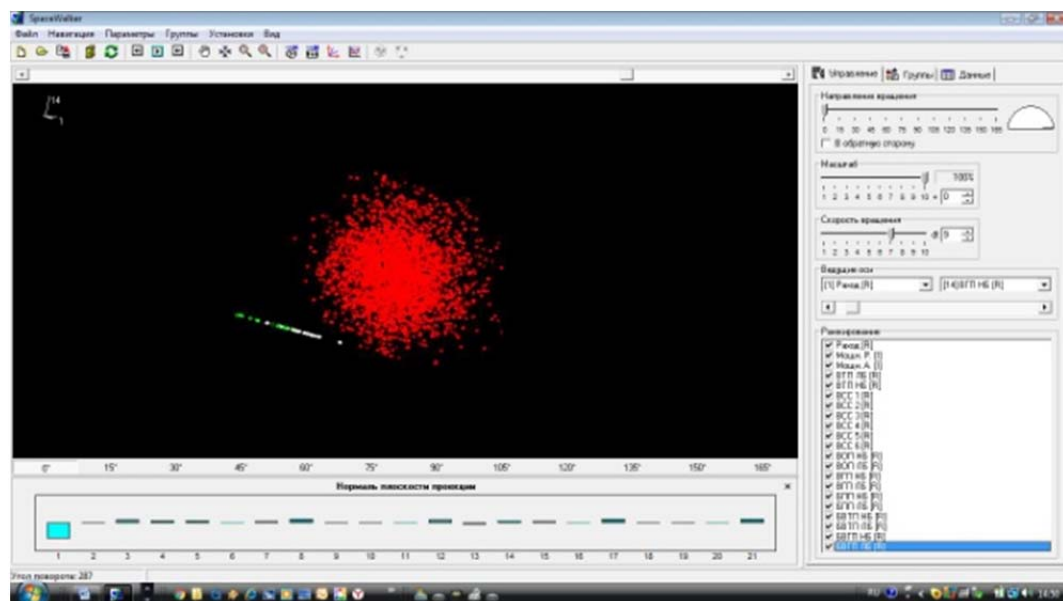


Рис. 17. Длительная временная зона аномально резкого изменения ряда параметров гидроагрегата

На обоих скриншотах стабильное во времени состояние комплекса параметров на когнитивном образе представляется в виде плотного облака, а временные тренды представляются в виде нитей и выбросов. Факт появления этих нитей и выбросов обусловлен выбираемыми проекциями параметров и соответственно индицирует как группу параметров обладающих аномальными свойствами, так и те временные интервалы, на которых эти аномалии появляются. Разумеется, количественное измерение связей между параметрами на данном временном интервале теперь становится возможным средствами многомерного статистического анализа.

6.6. Непараметрические статистики в многомерных временных рядах для когнитивного обнаружения аномалий

В настоящее время многомерный анализ временных рядов (который образуют изменяющиеся во времени характеристики объекта) получил широкое распространение в гидрологии, экологическом и медицинском мониторинге совокупности технических и природных объектов. Так, например, за годы наблюдений в гидрологической практике происходит накопление большого объема многомерных временных рядов (100 терабайт). Однако, их анализ представляет серьезные трудности (это и многомерность данных по характеристикам для одного технического объекта и многочисленность таких объектов). Кроме того, таблицы, описывающие эти объекты, содержат множество параметров, изменяющихся с течением времени по каждому объекту (скважине). Это так называемая вложенная иерархия многомерных данных [1]. Поэтому предлагается трех этапный способ вложенных многомерных массивов многомерных данных для выявления аномалий в поведении технических объектов.

6.6.1. Этапы когнитивного анализа

На первом этапе осуществляется когнитивный анализ [2–6] многомерных временных рядов в отдельных объектах с целью выявления характерных аномалий в поведении временных трендов (сигналов) по каждой характеристике объекта (например, скважине). На основании этого анализа можно сформировать модель сигналов в многомерных временных рядах и выделить фрагменты

рядов, не содержащих аномальных трендов со всей очевидностью. Это позволит использовать метод контраста при формировании модели сигнала (аномалии, тренда). Затем предлагаются непараметрические статистики количественной оценки этих аномалий, применяемые к многомерным временным рядам. На втором этапе с помощью этих статистик осуществляется анализ многомерных временных рядов для каждого объекта, что позволит представить каждый объект как точку в многомерном пространстве этих. Это своеобразное «свертка» многомерных рядов в набор точечных оценок величины аномалий в этих временных рядах параметров одного объекта. Оценки по отдельным параметрам объекта можно рассматривать как своеобразную нормировку временных рядов.

На третьем этапе весь набор объектов технического мониторинга можно представить как множество точек многомерном пространстве характеристик этих объектов. Это множество точек в многомерном пространстве характеристик можно снова представить виде когнитивного образа, теперь уже всего множества объектов. Только теперь выявление отдельных облаков в этом когнитивном образе будет показывать, насколько велики аномалии, и каким техническим объектам эти аномалии соответствуют. Другими словами можно рассматривать и наглядно видеть, как отличаются с точки зрения аномалий различные объекты (например, скважины) по всем многомерным характеристикам. По сути, предлагается метод, сочетающий объективную количественную робастную оценку многомерных аномалий с когнитивным наглядным образом множества скважин. Предлагается устойчиво и объективно обнаруживать аномалии в многомерных данных для одного объекта (здесь скважина), а затем с помощью когнитивных образов обобщать и наглядно выявлять эти многомерные аномалии для множества аналогичных объектов.

6.6.2. Устойчивые (робастные) алгоритмы выявления аномалий

Аномалии в поведении временных рядов могут быть описаны на основе следующей модели теории робастного обнаружения и оценивания сигналов [3]: $x(t)=s(t)+z(t)$. Здесь $x(t)$ это процесс на выходе регистрирующей аппаратуры представляет собой аддитивную смесь сигнала $s(t)$ неизвестной формы (для гидрологов аномалия) с шумом $z(t)$ (с неизвестным распределением, но описываемым параметрами сдвига, масштаба, независимости). По результатам наблюдений

формируем выборку $x_i = (x_1, \dots, x_i, \dots, x_k)$ и формируем выборку не содержащую сигнала (аномалии) $y_j = (y_1, \dots, y_2, \dots, y_m)$. Имея эти выборки можно методом контраста оценить энергию сигнала E_s (энергию аномалий). По аналогии с гауссовской моделью рассмотренной в [3] получаем оценку в виде: $E_s = M_e(x_{2i}) - M_e(y_{2j})$, здесь $M_e()$ – оценка медианы. Такие оценки получаются по всем характеристикам технического объекта (скважины) и в многомерном пространстве таких характеристик, каждый объект мониторинга представляет собой точку. Как отмечалось выше, множество объектов образует отдельные облака, которые характеризуют «похожесть» многомерных аномалий в разных облаках. Кроме того, уплощение облаков и их наклон к выделенным параметрам характеризует связь между аномалиями в терабайтных массивах данных. Еще одна модель сигнала из [3] может быть здесь использована. В случае отсутствия сигнала имеется выборка $x(t) = b + z(t)$, где b это неизвестная константа, $z(t)$ – случайная последовательность с неизвестным распределением, но описываемом параметрами сдвига, масштаба, независимости [3]. В случае наличия сигнала неизвестной формы $s(t)$ имеется выборка $y(t) = s(t) + b + z(t)$. По результатам наблюдений формируем выборку $y(t) = y_j = (y_1, \dots, y_2, \dots, y_m)$ и формируем выборку не содержащую сигнала (аномалий) $x(t) = x_i = (x_1, \dots, x_i, \dots, x_k)$. По аналогии с гауссовской моделью рассмотренной в [3] получаем статистику в виде: $\phi(x_i, y_j) = [M_e((y_j) - M_e(x_i))^2 / M_e((x_i) - M_e(x_i))^2]$. Эта квантильная статистика также позволяет выявлять наличие сигнала неизвестной формы даже при наличии неизвестной постоянной составляющей b . При этом эта статистика может выступать в качестве основы решающей функции $\phi(x_i, y_j) > C_\alpha$ для обнаружения сигнала и обладает соответствующими характеристиками качества обнаружения сигнала (ошибки первого α и второго рода $1 - \beta$), где: C_α пороговая константа.

6.6.3. Практическое применение

Предлагаемые методы ниже иллюстрируются на примере набора многомерных временных рядов для набора скважин месторождения минеральных вод «Нарзан». По каждой скважине имеются статистические данные, собранные за многие годы. В исследуемом многомерном массиве таблице в первом столбце представлены моменты времени от 1967 по 2013 года, в остальных столбцах представлены измеренные характеристики. Например, это $Na+K$, Ca , Mg , Cl , HCO_3 ,

SO₄, CO₂, сухой остаток, минерализация. В этом случае, каждый столбец будет представлять из себя изменяющиеся переменные во времени, т.е. временные ряды. Исходную таблицу можно представить так, будто заданы координаты для каждого момента времени в виде разных параметров по строке. Из этого следует, что каждый момент времени будет являться точкой в многомерном пространстве параметров. Для решения поставленной задачи использовалась программа когнитивной визуализации SpaceWalker [4]. Если открыть многомерные данные в программе, то на экране покажется некое облако точек, которое характеризует многомерный временной ряд. Если мы видим симметричное облако точек, то это означает, что с характеристиками скважины во времени ничего не происходит. Если изменения во времени есть, то будут наблюдаться изображения, в которых облако преобразуется в набор, напоминающий «горные хребты». Подобные изображения и позволяют визуализировать и анализировать многомерные временные ряды с целью предварительного обнаружения аномалий временных рядов.

Например, если выбрать такие ведущие оси Na+K и Mg, то выделится особая группа точек от основного облака (рис. 18).

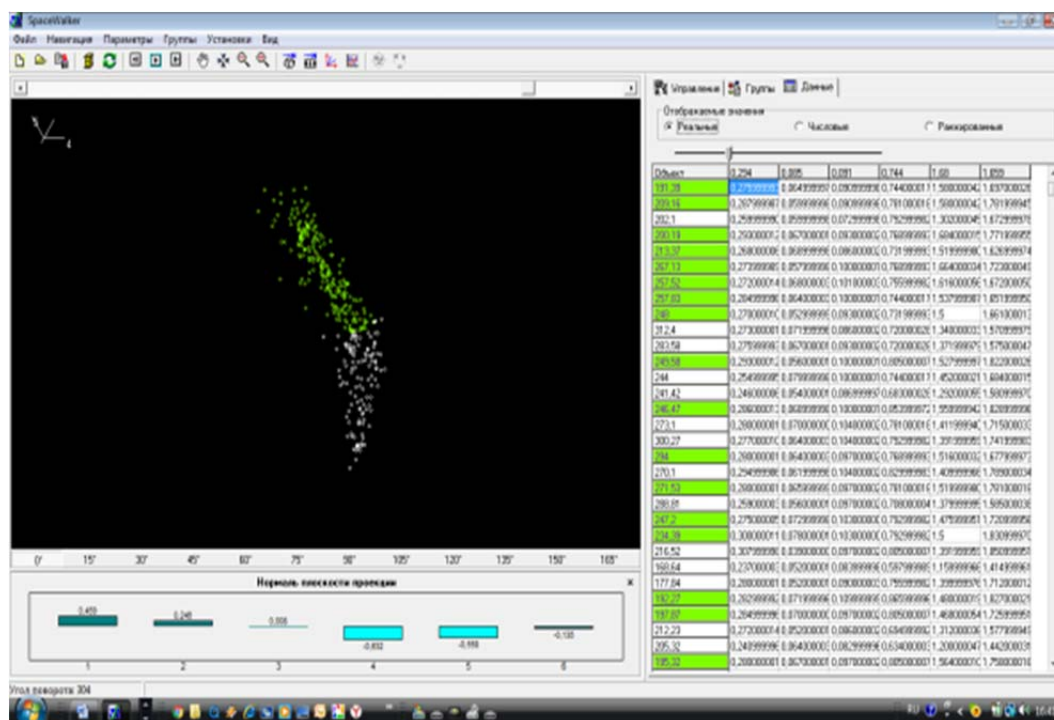


Рис. 18. Пример когнитивного выделения аномалий многомерных временных многомерных рядов для одной скважины

В исходной таблице эти точки относятся к промежутку времени от 1967 по 1974 года. Для выделенного промежутка времени характерно более низкое содержание CO_2 , HCO_3^- . Когнитивно и с помощью предложенных выше статистик можно выделить такие моменты и оценки размеров этих аномалий для набора скважин (рис. 19).

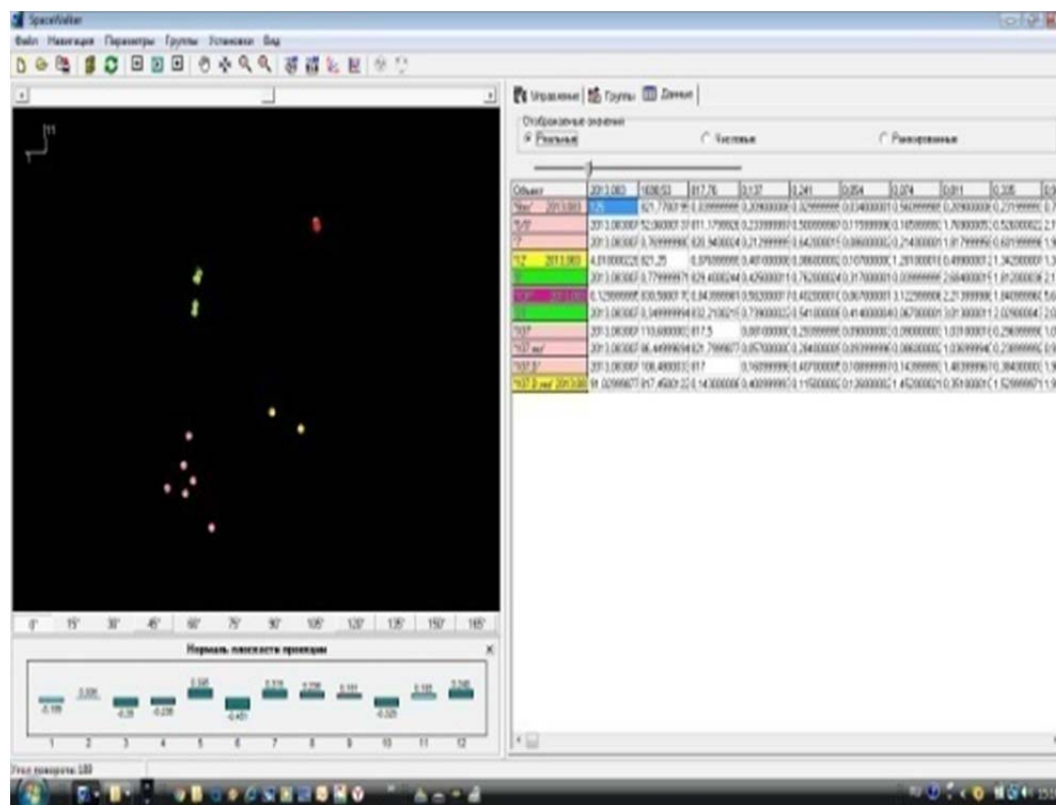


Рис. 19. Пример когнитивного робастного оценивания величины аномалий временных многомерных рядов для набора скважин

Заключение по п. 6.6

Предлагаемый когнитивный подход позволяет визуально заметить и количественно оценить малые, но многомерно связанные аномалии в терабайтных многомерных временных рядах для множества объектов мониторинга, что открывает новые эргономические возможности для мониторинга больших природных и техносферных систем. Современные системы статистического анализа, например пакет IBM SPSS 20 STATISTICS дополняют и обеспечивают детальный объективный количественный анализ обнаруженных с помощью когнитивных средств обнаружения техносферных и природных аномалий.

6.7. Практика формирования инженерных визуальных метафор на основе когнитивных образов многомерных данных техно-сферного мониторинга

Большой район КМВ относительно небогат ресурсами поверхностных вод. На его территории располагается часть бассейна реки Кумы с притоками, включая реку Подкумок, и реки Малки, входящей в бассейн Терека. Все реки района принадлежат бассейну Каспийского моря и берут начало на северном склоне Главного Кавказского и Скалистого хребтов, получая родниковое (Кума) или преимущественно ледниковое (Малка) питание. Для рационального использования минеральных вод осуществляется мониторинг ресурсов. Для этого измерялся широкий спектр гидрологических характеристик, из которых формировались многомерные массивы данных. Водоносные системы разной степени водообильности и площадного развития выделяют во всех стратиграфических комплексах пород. По концентрациям веществ широкого химического состава, минеральные воды весьма разнообразны и представлены гидрокарбонатной, сульфатной и хлоридной группами. Хлоридно-гидрокарбонатный натриевый, сульфатно-гидрокарбонатный кальциево-натриевый, гидрокарбонатно-хлоридный натриевый состав, в качестве специфических компонентов содержат углекислый газ, сероводород, йод и бром (Ессентуки), кремниевую кислоту (Железноводск).

Главное достоинство технологии когнитивного анализа представленных многомерных гидрологических данных, состоит в том, что не требуется построение предварительных гидрологических моделей [4]. Используя эту технологию для анализа, мы не вносим никаких искажений в сам процесс исследования. Все данные, с помощью алгоритмов когнитивного динамического проецирования представлялись в виде когнитивных образов. Визуальное наблюдение этих образов обеспечивало специалисту гидрологу возможность выявлять временные и пространственные экстремальные области с учетом более десяти гидрологических характеристик (рис. 20).

Для успешного практического использования когнитивных образов в настоящее время активно развиваются методы когнитивной психологии и гносеологии по выявлению и классификации когнитивных образов как визуальных инженерных метафор [5, 6]. Формируются онтологии подобных метафор позволяющие выявлять различные

аспекты экстремальных ситуаций. Так в данном случае срабатывает метафора «океанских волн» (рис. 21).

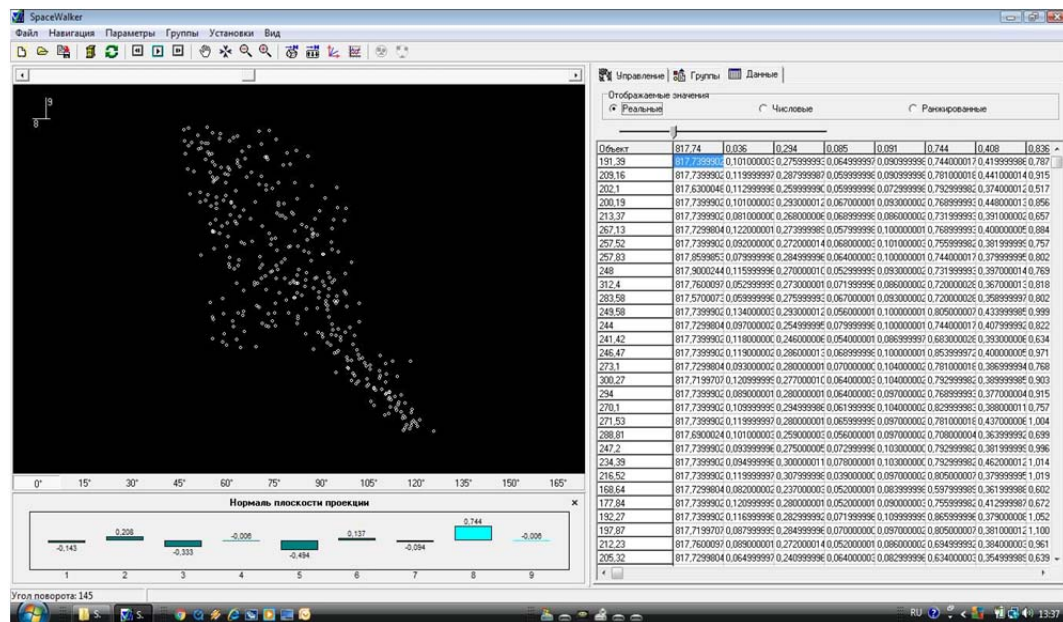


Рис. 20. Вид когнитивного образа многомерных временных рядов для типового поведения гидрологических параметров

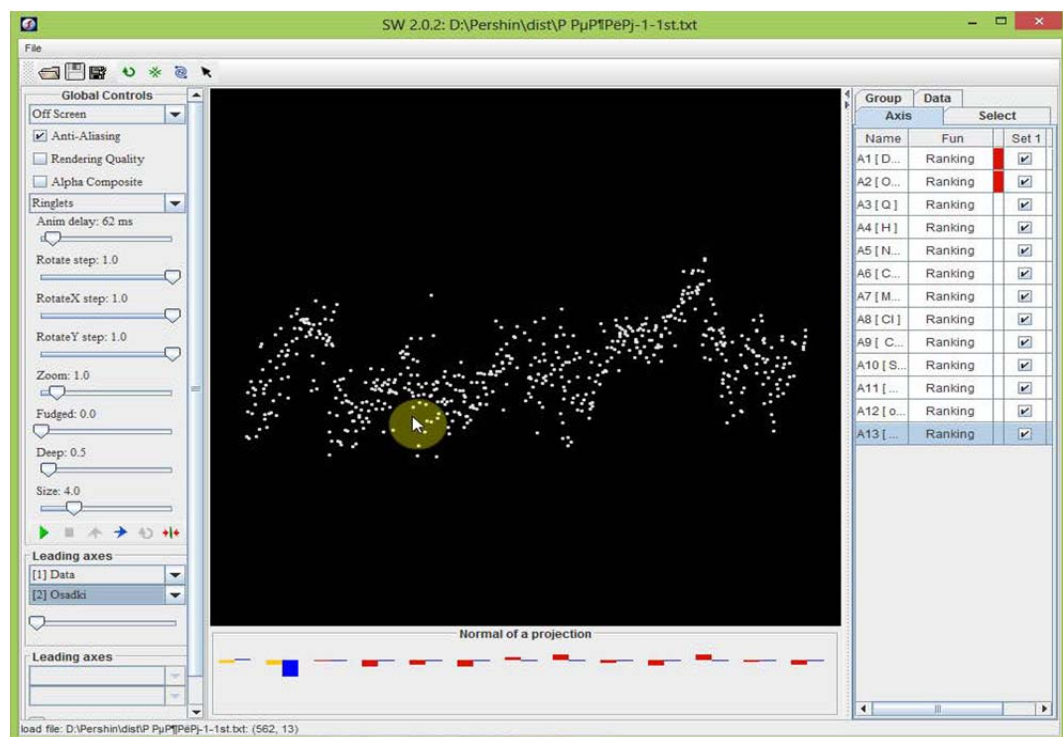


Рис. 21. Вид когнитивного образа многомерных временных рядов для нормальных периодических временных изменений гидрологических параметров

6.8. Когнитивные технологии визуализации многомерных данных для интеллектуальной поддержки принятия решений

В настоящее время приобретает актуальность проблема оперативного анализа большого объёма динамически изменяющихся параметров всего комплекса исследуемых объектов. Такая проблема возникает, например, в военной сфере при тактическом анализе боевых действий, техногенных катастроф, стратегическом планировании и моделировании использования комплексов вооружений. Она возникает и при создании нового поколения диспетчерских систем, отражающих обстановку в контролируемом воздушном или ином оперативном пространстве. Эти проблемы интенсивно разрабатываются как в рамках стратегического и тактического боевого искусства, с использованием всего арсенала современной математики (теория исследования операций, теория оптимального управления и оптимизации), так и в рамках создания автоматизированных комплексов современных вооружений [1].

Решение этих и других подобных проблем наталкивается на ряд существенных трудностей связанных с огромной ролью интуиции оператора, которая опирается на присущие человеку возможности непосредственного восприятия боевой обстановки или чрезвычайной ситуации. Современные условия боевых действий и техногенных катастроф оставляют оператора наедине с терминалами, где одновременно фиксируются тысячи параметров, которые человек не в состоянии оперативно воспринять и творчески «переработать» в своем сознании. Основная трудность состоит в том, что человек оказывается всего лишь элементом сложной автоматизированной системы контроля и управления, которая не адаптирована под творческие возможности человека. Разработанные ранее в рамках эргономики методы встраивания человека оператора в такую систему отчасти позволили адаптировать человека в так называемые эрготехнические системы, но огромный потенциал творческой и профессиональной интуиции не использовался в полной мере.

Однако сейчас, благодаря огромному прогрессу в области когнитивных наук, когнитивной психологии, гносеологии и информационных технологий [8, 9] появились принципиально новые возможности для радикального решения означенных выше проблем. Этот прогресс особо проявился в создании новых технологий и методик когнитивной машинной графики [2–5].

6.8.1. Принципы работы

Предлагаемый авторами подход позволяет осуществлять проекцию многомерных данных представленных в виде грассмановых многообразий на произвольно заданную оператором-исследователем плоскость в многомерном конфигурационном (фазовом) пространстве. При этом подбор наилучшего положения плоскости проекции осуществляет сам пользователь, опираясь на свою интуицию и когнитивный образ перед глазами. Имея возможность активно влиять на ориентацию плоскости проекции в многомерном пространстве, исследователь свободен от предварительных соображений о статистической (геометрической) структуре данных, которые представляют объекты. Человек непосредственно видит на экране проекции кластеров или многомерных поверхностей, в которые формируются его данные. Этот зрелищный образ стимулирует его интуитивное понимание исследуемых объектов [4–9].

Ниже дается краткий пример использования разработанных авторами средств когнитивной визуализации обстановки, способных решать поставленную выше задачу активной и управляемой стимуляции интуиции и эмпирического опыта оператора для принятия адекватных решений в современной сложной и быстроменяющейся обстановке. Кроме этого предлагаются и развиваются принципиально новые алгоритмические подходы, основанные на гиперболической геометрии и алгебраических многообразиях.

Пример когнитивной визуализации – когнитивный анализ техносферных опасностей, выполненный в рамках сотрудничества с МЧС России. Исследование проводилось с участием и экспертизой сотрудников ФГБУ ВНИИ ГОЧС (ФЦ) («Всероссийский научно-исследовательский институт по проблемам гражданской обороны и чрезвычайных ситуаций МЧС России» (Федеральный центр науки и высоких технологий)). В качестве исходных данных для анализа использована информация о чрезвычайных ситуациях, произошедших в 1 квартале 2012 года (703 чрезвычайных ситуации). Возникающие на сотнях объектов чрезвычайные ситуации (ЧС) анализировались по следующим выделенным параметрам:

- месяц;
- состояние;
- масштаб;
- регион;

количество пострадавших;
количество погибших;
личный состав;
техника;
источник ЧС.

Возможные варианты когнитивных образов в статическом положении для анализа этих ЧС (проекция многомерного облака на плоскость, заданную парой осей параметров) приведены ниже (рис. 22–24):

Можно сделать вывод, что использование визуализации многомерных статистических данных с помощью генерации когнитивного образа в качестве дополнительного инструмента при анализе и прогнозе чрезвычайных ситуаций позволило обратить внимание на особые классы ЧС, которые ранее, без использования интуитивного восприятия когнитивных образов обнаружить не удавалось.

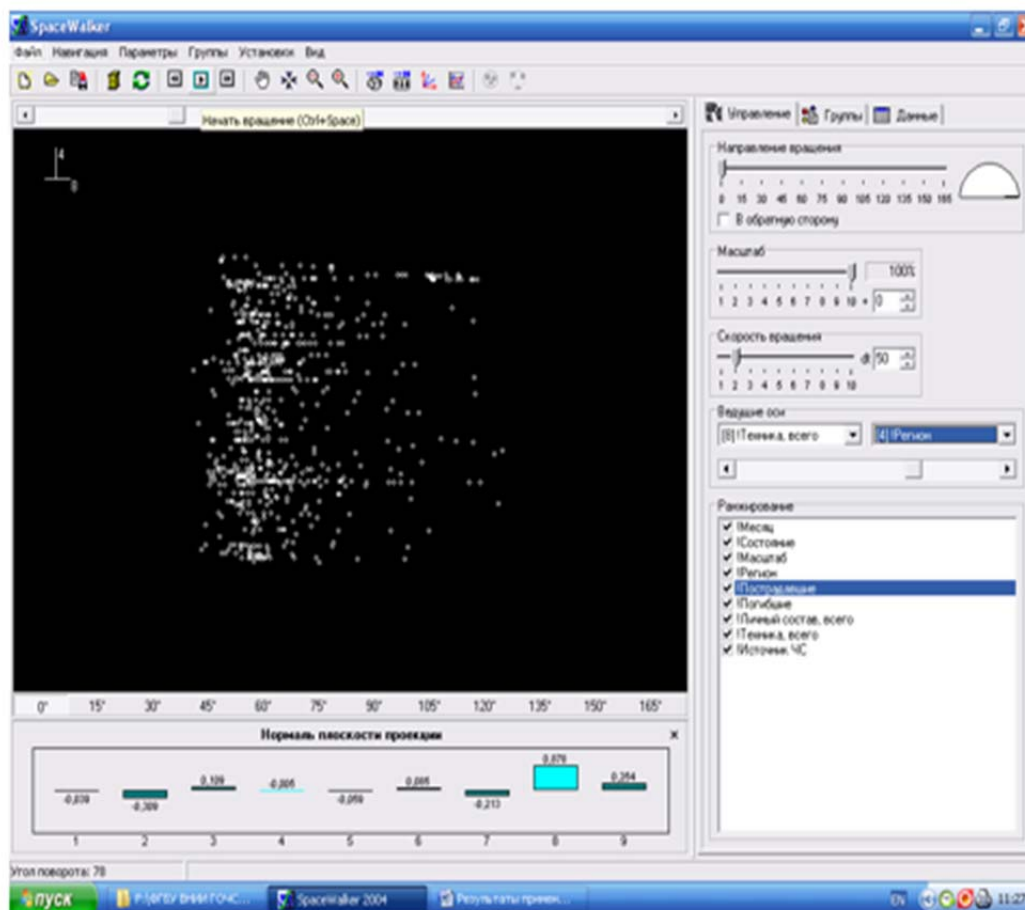


Рис. 22. Стратификация пострадавших при обеспеченности регионов техническими средствами спасения

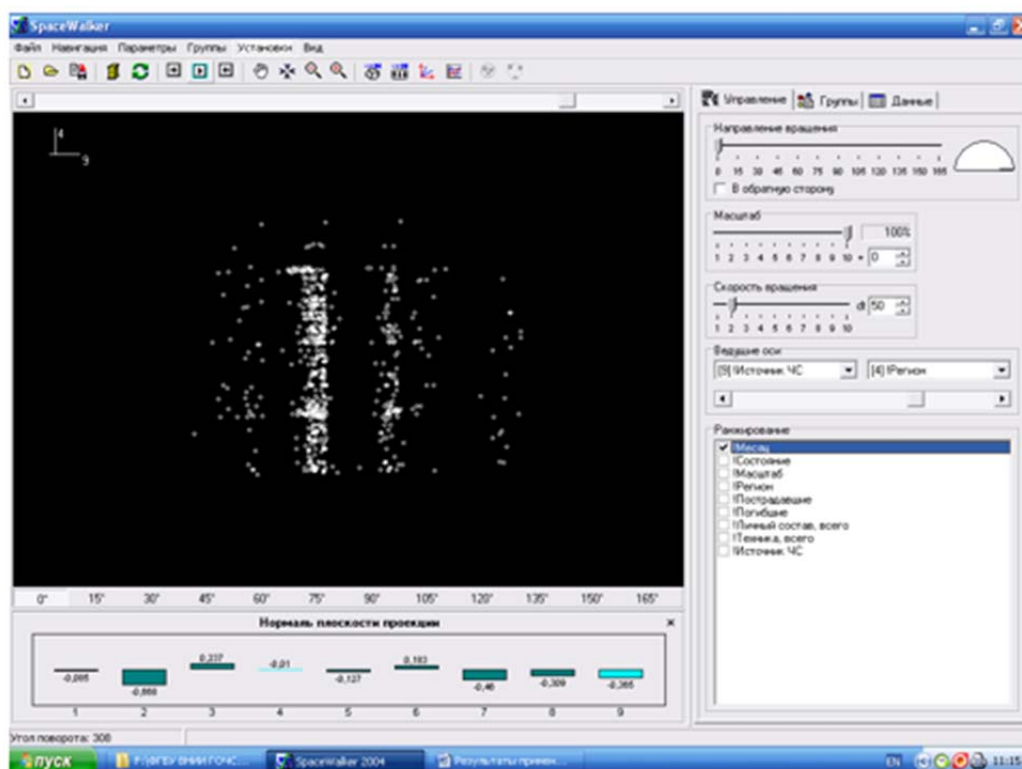


Рис. 23. Стратификация по срокам и регионам источников ЧС

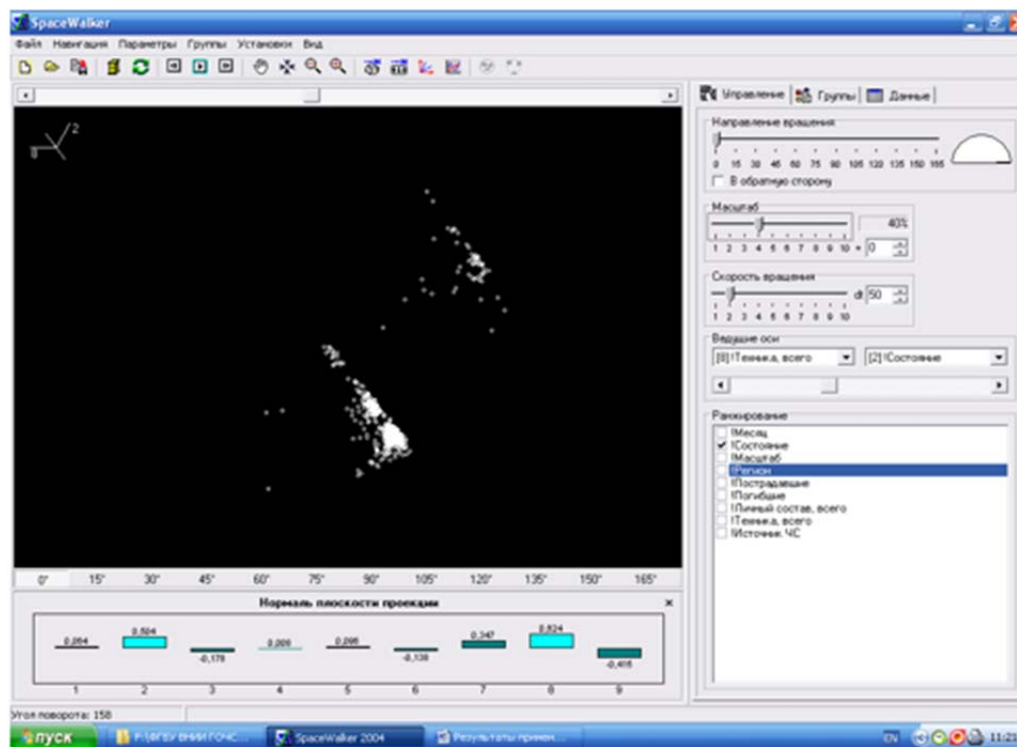


Рис. 24. Стратификация по регионам состояния и наличия технических средств спасения

6.8.2. Новые алгоритмы когнитивной визуализации

Предлагается дальнейшее развитие алгоритмов когнитивной визуализации на основе интерпретации k -мерного проективного пространства P_k в d -мерное гиперболическое пространство в kSd [6] с последующим преобразованием последнего в когнитивный псевдо-трехмерный образ. Такое формирование гиперболической геометрии многомерных данных происходит с использованием идеи плюккеровых координат [7]. Подобные алгоритмы способны когнитивно визуализировать даже терабайтные и даже петабайтные совокупности объектов.

Гиперболический алгоритм визуализации поддерживает эффективный режим взаимодействия с иерархиями гораздо большего размера, чем обычные средства визуализации иерархий. Если обычный двумерный визуализатор в окне 600х600 пикселей может отобразить 100 узлов, то гиперболический браузер может отобразить 1000 узлов, из них около 50 находятся в фокусе и легко читаемы.

Особенно это важно при анализе статистических связей, факторном анализе, обнаружении целей и при их распознавании. Тот факт, что процедура динамической визуализации не опирается на неполные и, возможно, ложные априорные сведения о природе объектов, а значит, не привнося в проекции искажающих влияний той или иной модели, дает возможность использовать визуализированные образы в условиях глубокой априорной неопределенности предметной области боевых действий и вооружений [1]. Авторами выполнена разработка мультиплатформенных Java-версий программных систем SpaceWalker и [10] способных реализовать технологии когнитивной визуализации оперативной обстановки для диспетчерских служб широкого профиля.

Здесь появляется еще одна возможность когнитивного контроля малейших изменений в состоянии объектов. Как показали исследования, даже малые изменения параметров объектов существенно изменяют их когнитивные образы, что позволяет человеку-оператору мгновенно заметить изменение характеристик объектов. Еще раз подчеркнем, что использование многомерной гиперболической геометрии при создании когнитивного образа позволяет визуально представлять содержание терабайтных многомерных массивов. Кроме того, использование перечисленных применений когнитивной графики будет еще более эффективным при ее внедрении в сетевые технологии. Впечатляющий эффект может быть получен при внедрении метода оперативного анализа в онлайн-системах космического мониторинга [10].

Выводы по п. 6.8

Можно быть уверенным, что разработанные методики, алгоритмы и прикладные когнитивные программные системы найдут широкое применение в различных отраслях, где необходим оперативный анализ большого объема многомерных данных – от планирования операций до мониторинга и моделирования технических систем.

6.9. Семиотика визуальных инженерных метафор когнитивных образов многомерных данных – новая парадигма компьютерной эргономики

Идея когнитивной визуализации многомерных данных привела к созданию ряда компьютерных инструментов SW-15, Space Hedehog, GGOBI порождающих когнитивные образы в сознании пользователя [1, 2]. Генерация практически полезных когнитивных образов, обеспечивших поддержку принятия решений при обработке DIG DATA в условиях априорной неопределенности поведения сложных систем в свою очередь создала проблему, создания методик для эффективного использования этого нового феномена нашего сознания.

Прежде всего, необходимо использовать достижения биологов (Умберто Мартурано) и гносеологов (Девида Чалмерса, Д. Денета), которые показали, что механизмы человеческого познания обладают огромным неожиданно открытым потенциалом, обеспечивающим устойчивое принятие решений в условиях априорной неопределенности окружающей среды [3, 4, 5].

Кроме этого при попытках использовать возможности этого потенциала человеческого сознания в процессе проектирования программных систем обработки многомерных данных инженеры программисты представили многомерные данные как визуальные метафоры а инструменты их обработки как процедуры манипуляции над ними [5, 6, 7]. Таким образом, все аспекты программирования идентифицируются как графические модели (графические символы, иконы) предстающими для человеческого сознания, как визуальные метафоры. Визуальные метафоры являются феноменами, которые когнитивные психологи выявляют как продукт деятельности метафорического мышления. Напомним, что это один из механизмов познания наряду с другими механизмами, например – рациональное

мышление. Сама идея воспользоваться метафорическим мышлением для принятия решений при обработке многомерных данных петабайтного масштаба (когда традиционные методы рационального мышления дают сбой) необычайно ценна, полезна и богата своими возможностями управления интуитивными возможностями нашего сознания [6, 7].

В данной работе как показано выше на основе достаточно значительного набора решенных предметных задач анализа многомерных данных (наблюдательная астрофизика, мониторинг подвижного состава РЖД, мониторинг ВУЗОВ, мониторинг объектов здравоохранения, мониторинг гидрогеологии источников минеральных вод) предлагаются методы классификации полученных когнитивных образов как онтологий визуальных метафор.

Кроме того появляется возможность нового использования методики объективации когнитивных образов и визуальных метафор на основе систем Flash MX и 3D MAX для объективации наблюдаемых когнитивных образов в виде стандартных 3D графических моделей. Примеры характерных визуальных метафор представлены на рис. 25–27.

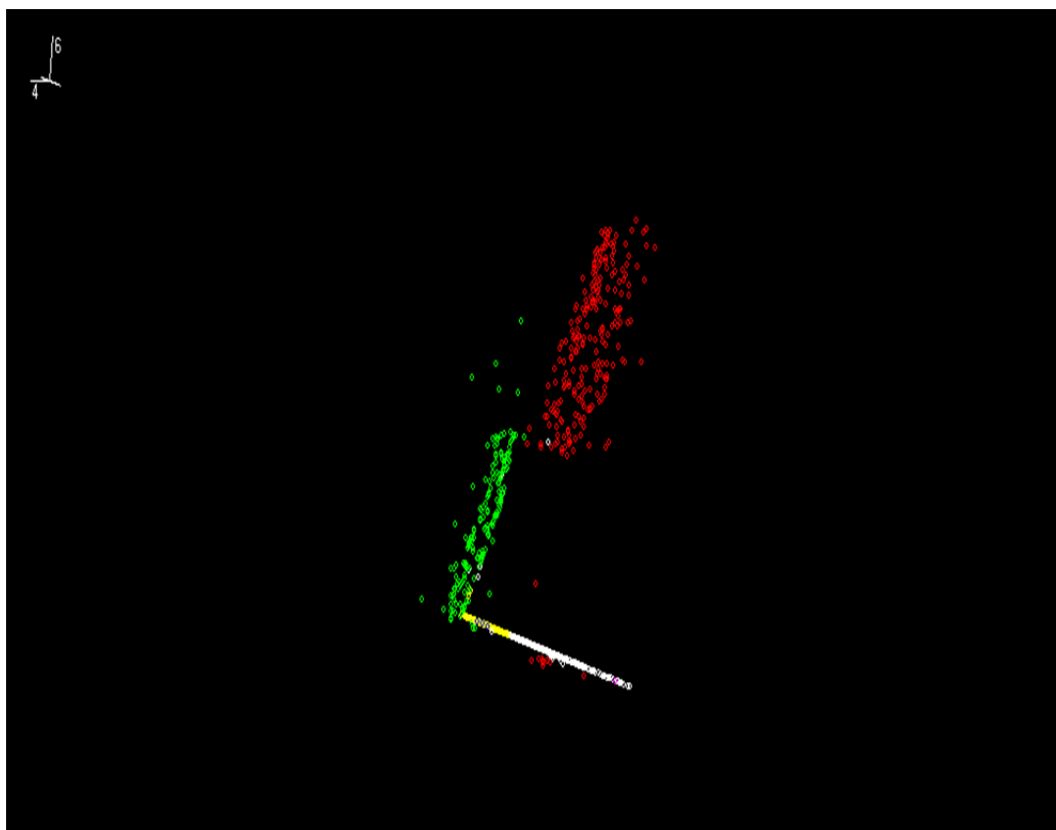


Рис. 25. Метафора «Трамплин»

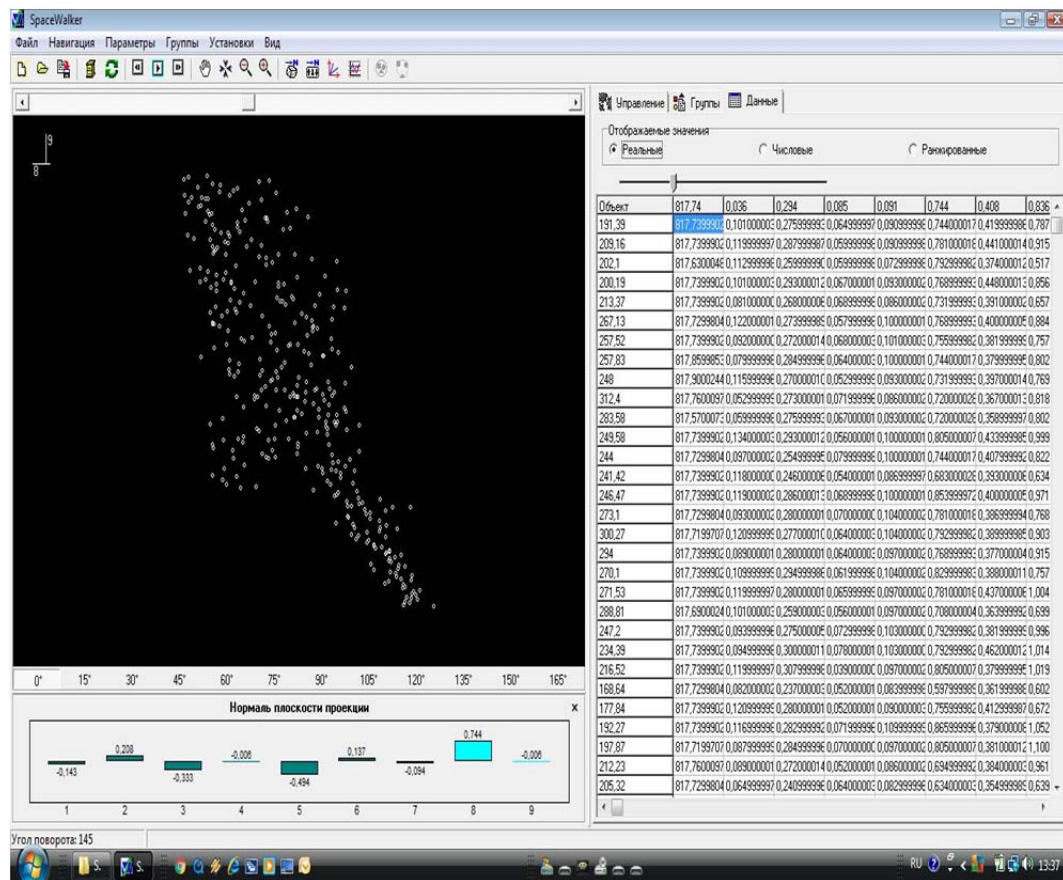


Рис. 26. Метафора «Бутылка Клейна»

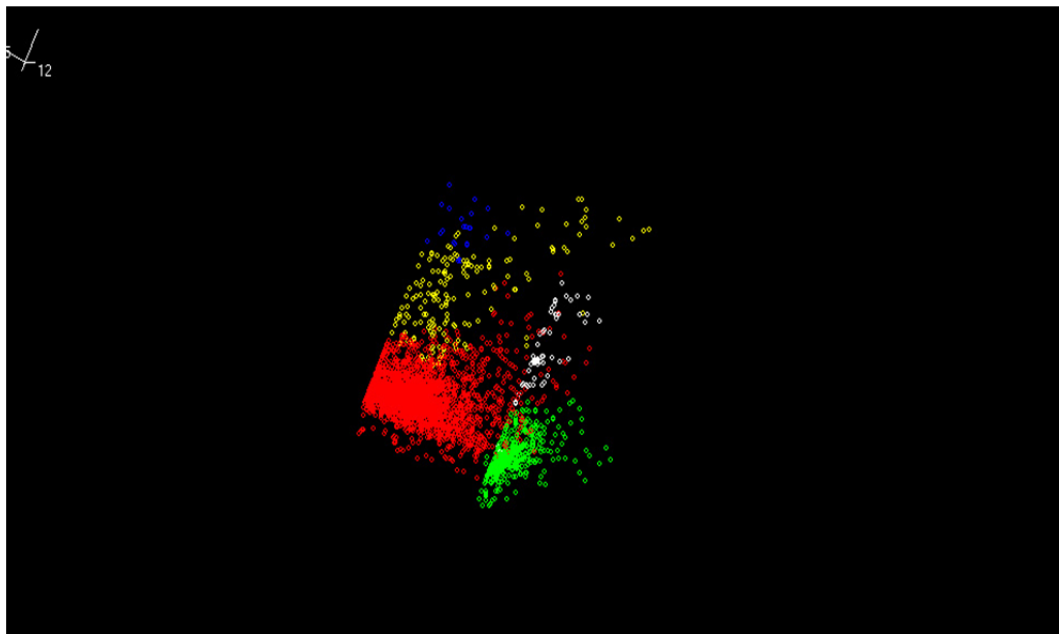


Рис. 27. Метафора «Раскрытая книга»

Заклучение

В настоящее время приходит понимание того факта что развитие и эволюция техносферы будет определяться процессами углубления и усиления взаимодействия машины с интеллектом человека-оператора в рамках теории автопоэзиса (Ф. Варелла, У. Матурана). Это приводит к усилению роли эргономики как научного направления, которое отвечает за интеграцию ресурсов человеческого сознания и техносферы. По сути происходит переход от парадигмы «адаптация машин – под физиологию человека» к парадигме «симбиоз интеллекта человека и машины» и наконец переход к формированию техносферы на новом уровне, когда машина и человек сливаются в новое образование имеющее принципиально новую онтологию и структуру. Речь фактически идет о метаморфозе, в чем то напоминающем метаморфоз в биологии.

Ситуация стимулируется бурным прогрессом в области искусственного интеллекта, аналитической философии, нейрофизиологии, когнитивной психологии.

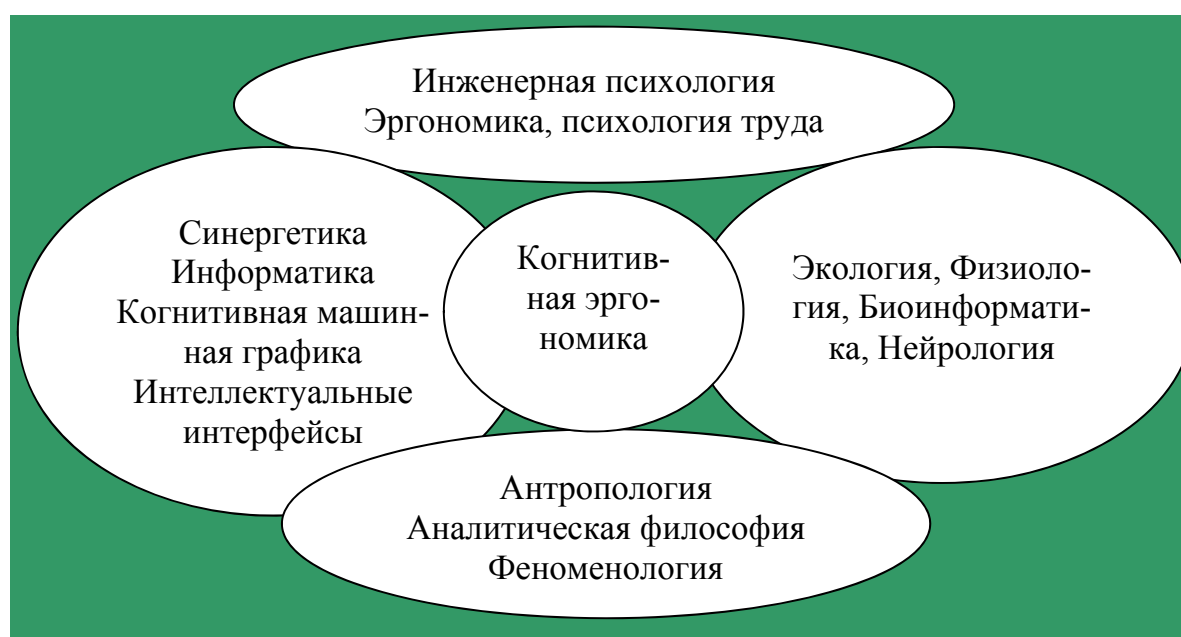


Рис. 28. Взаимосвязь когнитивной эргономики с психологией, философией, экологией и информатикой

Таким образом, все эти направления развития современной науки пересекаются между собой, образуя когнитивную эргономику и возможно способствуя формированию мощных средств обеспечения

техносферной безопасности. Информационные технологии впитывают перечисленные выше достижения и воплощают их в виде интеллектуальных интерфейсов и когнитивной машинной графики. Когнитивная эргономика – новое направление исследований эрготехнических систем. Наглядно это представлено на рис. 28.

Список литературы

Литература по п. 6.2

1. *Andrew Lo*. Big data, Systemic Risk, And Privacy-Preserving Risk Measurement / Big Data & Privacy – Work Shop Summary Report June 19, 2013 Massachusetts Institute of Technology 2013.
2. Проблемы сознания в философии и науке / под ред. проф. Д.И. Дубровского. М.: «КАНОН+» РООИ «Реабилитация», 2009. 427 с.
3. Когнитивный подход / отв. ред. – академик РАН В.А. Лекторский. М.: «КАНОН+» РООИ «Реабилитация», 2008. 464 с.
4. *Горохов В.Л., Лукьянец А.А., Чернов А.Г.* Современные методы когнитивной визуализации многомерных данных. Монография. Томск.: Некоммерческий фонд развития региональной энергетики, 2007. 216 с.
5. *Горохов В.Л., Тееррикортн П., Барышев Ю.В., Муравьев И.П.* Компьютерные метафоры феноменологических конструкций Ф. Брентано М. Хайдегера Ж.П. Сартра для систем обработки астрофизических наблюдений / Труды международной конференции по мягким вычислениям, 25–27 июня 2007, (SCM – 2007). Санкт-Петербург, 2007.
6. *Vitkovskiy V., Gorohov V., Komarinskiy S.* 6-D Visualization of Multidimensional Data by Means of Cognitive Technology. Astronomical Data Analysis Software and Systems (ADASS) XIX / Editors: Mizumoto Y., Morita K.-I., Ohishi M.; San Francisco, USA, 2010. Pp. 449–553.
7. *Широков С.И., Теханович Д.И., Барышев Ю.В.* Флюктуации пространственного распределения галактик в глубоком обзоре COSMOS на масштабах гигапарсеки. Вестн. С.-Петербург. Ун-та Сер 1. 2014.

Литература по п. 6.3

1. *Гаррет Р., Лондон Дж.* Основы операций на море М.: Военное издательство М.О., 1974. 269 с.
2. *Cook D, Swaine D.E.* Interactive and Dynamic Graphics For Data Anlysis. Spriger. 2009. 345 с.

3. *Горохов В.Л., Муравьев И.П.* Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Монография / под ред. проф. А.И. Михайлушкина. СПб. СПбГИЭУ, 2008. 172 с.

4. *Горохов В.В., Цаплин В.В.* Когнитивные технологии визуализации многомерных данных для интеллектуальной поддержки принятия решений / Цаплин В.В. Горохов В.Л., Витковский В.В. // Программные продукты и системы. 2014. № 3(107). С. 22–25.

5. *Горохов В.Л., Цаплин В.В., Витковский В.В.* Разработка когнитивной технологии оперативного управления сложными системами / Труды XV Всероссийской научно практической конференции РА РАН «Актуальные проблемы защиты и безопасности» Военно-морской флот России. 3–6 апреля 2012 г. СПб.: НПО Спецматериалов. Т. 4. С. 129–132.

Литература по п. 6.4

1. *Vitkovskiy V.V., Gorochov V.L.* New method of data mining in practical cosmology / Proceedings International conference «Problems of Practical Cosmology» Saint-Petersburg 2008. Pp. 256–259. (http://ppc08.astro.spbu.ru/fin_program.html).

2. *Gorokhov V.L., Evdokimov V.A., Vitkovskiy V.V.* Cognitive multidimensional data visualization in analyzing and decision_support systems / Третья Международная конференция по когнитивной науке 20–25 июня 2008 г., Москва, Россия, Тезисы докладов том 1 (The Third international conference on cognitive science. June 20–25, 2008, Moscow, Russia abstracts volume 1). Москва, 2008. С. 55.

3. *Komarinskiy S., Vitkovskiy V., Gorohov V., Zakharevski D.* The Cognitive Visualization System Astronomical Data Analysis Software and Systems (ADASS) XVII / Editors: Argyle, Robert W.; Bunclark, Peter S.; Lewis, James R. London, UK 2008. 394 p.

4. *Горохов В.Л., Тееррикори П., Барышев Ю.В., Муравьев И.П.* Компьютерные метафоры феноменологических конструкций Ф. Брентано, М. Хайдегера Ж.П. Сартра для систем обработки астрофизических наблюдений / Труды международной конференции по мягким вычислениям, 25–27 июня 2009, (SCM – 2009). Санкт-Петербург, 2009.

5. *Горохов В.Л., Барышев Ю.В., Тееррикори П., Витковский В.В.* Когнитивная визуализация космологических обзоров без привязки к сетке космологических моделей / Сборник научных трудов научно-

практической конференции по современным проблемам прикладной информатики» 19–20 мая 2009 г. СПб.: Издательство СПбГПУ, 2009. С. 146–149.

6. Струков Д.Р., Горохов В.Л. Геоинформационные системы и многомерные статистические методы пространственного анализа // Информационно-управляющие системы. Т. 3(40). 2009. С. 56–62.

7. Gorokhov V. Cognitive imaginig in visual data-driven decision-support systems / Proceedings International conference «Astronomical Data Analysis Software and systems». Sapporo. Japan. 4–8 October 2009. Pp. 5.

8. Evdocimov V., Gorokhov V., Vitkovskiy V. Cognitive imaginig in visual data / Abstract. Scientific programme. International conference «High Performance Computing in Observational Astronomy: Requirements and Challenges». Puna. India. 12–14 October. 2009. Pp. 5.

9. Горохов В.Л., Террикорни П., Барышев Ю.В. Конструирование компьютерных метафор на основе феноменологических онтологий для систем когнитивной машинной графики / Сборник научных трудов. СПб.: Издательство СПбГИЭУ, 2009. С. 29–54.

Литература по п. 6.5

1. The United Kingdom Offshore Oil and Gas Industry Association (Oil and Gas UK) / Accident atistics for Offshore Units on the UKCS 1990–2007, 2009. URL: <http://www.oilandgasuk.co.uk/cmsfiles/modules/publications/pdfs/EC024.pdf>.

2. Горохов В.Л., Муравьев И.П. Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Методология, методики и интерфейсы. Монография. СПб.: ИНЖЭКОН, 2007. 173 с.

Литература по п. 6.6

1. Чесалов С.М. Статистические методы решения гидро-геологических задач на ЭВМ / Чесалов С.М. Москва: Изд-во Недра, 1989. 178 с.

2. Прокопчина С.В. Управление в условиях неопределенности / Шестопапов М.Ю., Уткин Л.В Лазарев В.Л. СПб.: Изд-во СПбГЭТУ, 2014. 295 с.

3. Богданович В.А. Теория устойчивого обнаружения, различения и оценивания сигналов / Богданович В.А., Вострецов А.Г. М.: Изд-во Физматлит, 2003. 316 с.

4. *Горохов В.Л.* Современные методы когнитивной визуализации многомерных данных / Горохов В.Л., А.А. Лукьянец, А.Г. Чернов монография. Томск.: Из-во Некоммерческий фонд развития энергетики, 2007. 216 с.

5. *Чалмерс Д.* Сознательный ум. В поисках фундаментальной теории / Монография. М.: Изд-во URSS, 2014. 509 с.

6. *Зенкин А.А.* Когнитивная компьютерная графика / Зенкин А.А. Под ред. Д.А. Поспелова. М.: Изд.-во Наука, 1991. 192 с.

Литература по п. 6.7

1. Когнитивный подход / отв. ред. – академик РАН В.А. Лекторский. М.: «КАНОН+» РООИ «Реабилитация», 2008. 464 с.

2. *МакКормак Э.* Когнитивная теория метафоры. / пер. с англ. // Теория метафоры. М.: Прогресс, 1990. С. 358–387.

3. *Cook D., Swaine D.E.* Interactive and Dynamic Graphics For Data Anlysis. Spriger. 2009. 345 с.

4. *Горохов В.Л., Муравьев И.П.* Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Монография. Под ред. проф. А.И. Михайлушкина. СПб.: СПбГИЭУ, 2008. 172 с.

5. *Авербух В.Л.* К теории компьютерной визуализации // Вычислительные технологии. 2005. Том 10. № 4. С. 21–41.

5. *Блэк М.* Метафора // Теория метафоры. М., 1990.

Литература по п. 6.8

1. *Гаррет Р., Лондон Дж.* Основы операций на море. М.: Военное издательство М.О., 1974. 268 с.

2. *Зенкин А.А.* Когнитивная компьютерная графика. М.: Наука. 1991.

3. *Cook D, Swaine D.E.* Interactive and Dynamic Graphics for Data Anlysis. Spriger. 2009. 345 с.

4. *Горохов В.Л., Муравьев И.П.* Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Монография. Под ред. проф. А.И. Михайлушкина. СПб. СПбГИЭУ, 2007. 170 с.

5. *Andrew Lo.* Big data, Systemic Risk, And Privacy-Preserving Risk Measurement / Big Data & Privacy – Work Shop Summary Report June 19, 2013 Massachusetts Institute of Technology, 2013. 45 p.

6. Розенфельд Б.А. Многомерные пространства. М.: Наука, 1966. 647 с.
7. Клейн Ф. Высшая геометрия. М.: УРСС, 2004. 400 с.
8. Когнитивный подход / отв. ред. – академик РАН В.А. Лекторский. М.: «КАНОН+» РООИ «Реабилитация», 2008. 464 с.
9. Управление в условиях неопределенности / Прокопчина С.В., Шестопалов М.Ю., Уткин Л.В., Куприянов М.С., Лазарев В.Л., Имаев Д.Х., Горохов В.Л., Жук Ю.А., Спесивцев А.В. Монография. СПб.: Из-во: СПбГЭТУ «ЛЭТИ», 2014. 303 с.
10. Vitkovskiy V., Komarinskiy S. 6-D Visualization of Multidimensional Data by Means of Cognitive Technology. Astronomical Data Analysis Software and Systems (ADASS) XIX / Editors: Mizumoto Y. Morita K.-I., Ohishi M.; San Francisco, USA, 2010. Pp. 449–553/

Литература по п. 6.9

1. Cook D., Swaine D.E. Interactive and Dynamic Graphics For Data Analysis. Spriger. 2009. 345 с.
2. Горохов В.Л., Муравьев И.П. Когнитивная машинная графика. Методы динамических проекций и робастная сегментация многомерных данных. Монография. Под ред. проф. А.И. Михайлушкина. СПб.: СПбГИЭУ, 2008. 172 с.
3. Дэвид Чалмерс. Сознательный ум. М.: URSS, 2013. 498 с.
4. Франсиско Х. Варела, Умберто Р. Матурана «Древо познания» М.: URSS, 2003. 252 с.
5. Д. Денет. Сладкие грезы. Чем философия мешает науке о сознании. М.: URSS, 2016. 292 с.
6. Горохов В.Л. Управление в условиях неопределенности. / Прокопчина С.В., Шестопалов М.Ю., Уткин Л.В., Куприянов М.С., Лазарев В.Л., Имаев Д.Х., Горохов В.Л., Жук Ю.А., Спесивцев А.В. Монография. СПб. Из-во СПбГЭТУ «ЛЭТИ», 2014. 303 с.
7. Мак Кормак Э. Когнитивная теория метафоры. / Пер. с англ. // Теория метафоры. М.: Прогресс, 1990. С. 358–387.

Литература по заключению

1. Матурана У.Р., Варела Ф.Х. Древо познания. Биологические корни человеческого понимания. М.: «Прогресс-Традиция», 2001.
2. Сергеев С.Ф. Обучающие и профессиональные иммерсивные среды / С.Ф. Сергеев. М.: Народное образование, 2009. 432 с.

3. *Cook D., Swaine D.E.* Interactive and Dynamic Graphics for Data Anlysis. Spriger. 2009.

4. Проблемы сознания в философии и науке / под ред. проф. Д.И. Дубровского. М.: «КАНОН+» РООИ «Реабилитация», 2009. 427 с.

5. *Чалмерс Д.* Сознający ум: В поисках фундаментальной теории. М.: URSS Книжный дом «ЛИБРОКОМ», 2013. 512 с.

6. Когнитивный подход / отв. ред. – академик РАН В.А. Лекторский. М.: «КАНОН+» РООИ «Реабилитация», 2008. 464 с.

7. *Горохов В.Л., Лукьянец А.А., Чернов А.Г.* Современные методы когнитивной визуализации многомерных данных. Монография. Томск.: Некоммерческий фонд развития региональной энергетики, 2007.

8. *Горохов В.Л.* Быть или не быть: проблема мироздания в современной науке и технике / Семенов А., Малышев Ю., Горохов В. / Монография. Palmarium Academic Pulishing. Ist ein Imprint der / OmniS-criptun GmbH & Co. KG Heinrich-Bocking-Str. 6-8, 66121 Saarbrucken, Deutchland, 2015. Pp. 97.

ГЛАВА 7

НЕЧЕТКИЕ МЕТОДЫ САМООРГАНИЗАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ

М.С. Куприянов, Л.Г. Логинская

Механизм самоорганизации является мощным инструментом адаптации структуры и поведения информационной системы к изменению внутреннего состояния и внешних условий. В общем случае механизм восстановления работоспособности должен обеспечивать:

- фиксацию событий, возникающих в системе;
- измерение качества функционирования отдельных блоков и системы в целом в процессе функционирования (например, динамическое измерение работоспособности с учетом перемежающихся отказов);
- сравнение оценок работоспособности (в общем случае качественных);
- принятие решений в условиях неполной и неточной (качественной) информации о типе архитектуры системы управления и ее составе с учетом заданного критерия оптимизации (наибольшей производительности или длительности функционирования, готовности к реализации выделенного набора функций и т.п.);
- принятие решений о наборе основных и вспомогательных функций, обязательных к реализации на рассматриваемом этапе функционирования системы;
- при снижении работоспособности всех компонентов системы (в том числе и резервных) перераспределение функций, либо переход к функциям низшего ранга.

Теоретически доказано, что мощность топологического пространства выше мощности метрического пространства [1], что определяет необходимость создания для объектов с неопределенными параметрами и свойствами адаптивных качественных шкал.

В данном разделе рассматриваются:

- метод переградуирования качественных шкал для реализации механизмов самоорганизации;
- метод динамического измерения работоспособности самовосстанавливающихся информационных систем (СИС).

7.1. Нечеткие числа и логические значения

Нечеткие числа представляют собой некоторые области базовых значений Q с определенными на них колокообразными функциями $Q \rightarrow (0, 1)$, (эти функции называются размытием нечеткого числа) и с выделенным базовым значением, соответствующем максимуму функции – размытия (это значение называется ядром или «модой» нечеткого числа). Максимум размытия меньше или равен 1, по определению.

Нечеткие числа, в зависимости оттого, что является базовой областью значений и какие операции определены над ними, называются арифметическими или логическими.

Арифметические нечеткие числа определены на всей вещественной оси $(-\infty, +\infty)$, над ними определяются операции арифметического сложения, умножения, вычитания, деления. Логические нечеткие числа определены на отрезке $(0, 1)$, над ними определяются логические операции: конъюнкция, дизъюнкция, отрицание. Кроме того, над обоими этими типами нечетких чисел определяется операция модификации, связанная с изменением формы размытия (возможно, сопровождающейся изменением самого ядра нечеткого числа).

Нечеткие числовые переменные и операции получения нечетких чисел из обычных чисел R_1 .

Будем использовать следующие обозначения: A – произвольное нечеткое множество над R_1 , $T_a(x)$ – функция принадлежности нечеткого множества A .

Определение 1. Назовем A нечетким числом, если $T_a(x)$ имеет вид одnogорбой функции, т.е. функции с единственным максимумом, левая ветвь которой неубывающая функция, а правая – невозрастающая.

Определение 2. Назовем ядром нечеткого числа A отрезок $J(R_1)$, на котором функция $T_a(x)$ достигает максимума: ядро $(A) ::= \{x \mid \forall x \in R_1, T_a(x \in R_1) = \max T_a(x)\}$. Размытием A будем называть саму функцию $T_a(x)$.

Определение 3. Нечеткое число A назовем простым нечетким числом, если ядро A представляет из себя единственную точку $a \in R_1$. В дальнейшем мы будем использовать название «терм» для обозначения простого нечеткого числа.

Произвольные нечеткие множества над R_1 (нечеткие числовые множества) можно представить как результат объединения нечетких чисел, ядрами которых являются точки локальных максимумов, а размытия определены таким образом, чтобы в результате объединения чисел как множеств получалось исходное нечеткое множество.

Определение 4. Простые нечеткие числа получаются как обобщение обычных чисел при помощи операции увеличения нечеткости.

Пусть $X \in R_1$ – обычное число, тогда $H(X)$ – нечеткое число, где H – операция увеличения нечеткости (операция расплывания).

Определение 5. В языке существуют термины, соответствующие операции расплывания, например: «почти», «немного больше», «немного меньше» и т.п. Назовем эти операции базовыми операциями нечеткости. Определим числа, соответствующие этим операциям, для базового ядра 0 (при помощи опроса экспертов, например). Назовем их базовыми «почти – числами». На основе этих чисел определим подобные числа с другими ядрами. Все такие числа для определенности будем называть «почти – числами».

Принцип построения «почти – чисел»: размытие чисел, «почти» совпадающих с «а», должно увеличиваться при $a \rightarrow \infty$. Например, «почти 1" это где-то от 0 до 2, а «почти 1000" находится где-то от 900 до 1100. Примем обозначение: «почти – числа» будем обозначать тремя символами – первый – T , терм; второй – p , «почти – число»; третий – значение ядра этого числа. Например, T_{pa} означает терм «почти – числа» с ядром в точке a .

Определение 6. Нечеткие «почти – числа» определим на основе базовых «почти – чисел». Пусть T_{p0} – базовое «почти – число». Тогда $T_{pa} = H_a(T_{p0})$, где H_a – оператор изменения размытия базового числа, зависящий от a .

Например, множество чисел для термина «немного больше». Допустим, базовый терм «немного больше 0" задан так, как на рис. 1. Тогда терм «немного больше а» может выглядеть так, как на рис. 2.

В общем случае, вид функций принадлежности зависит не только от принципа сформулированного выше, но и от того, какие еще термы определенной шкалы используются при принятии решений, т.е. совместный анализ условий влияет на вид функции принадлежности.

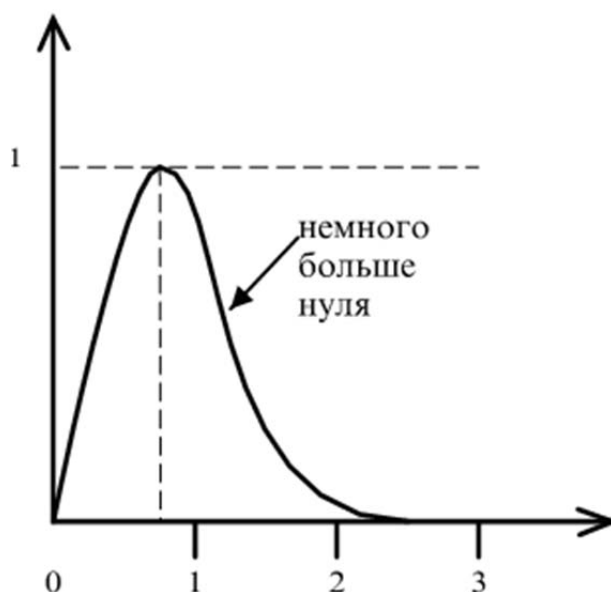


Рис. 1. Терм «немного больше 0»

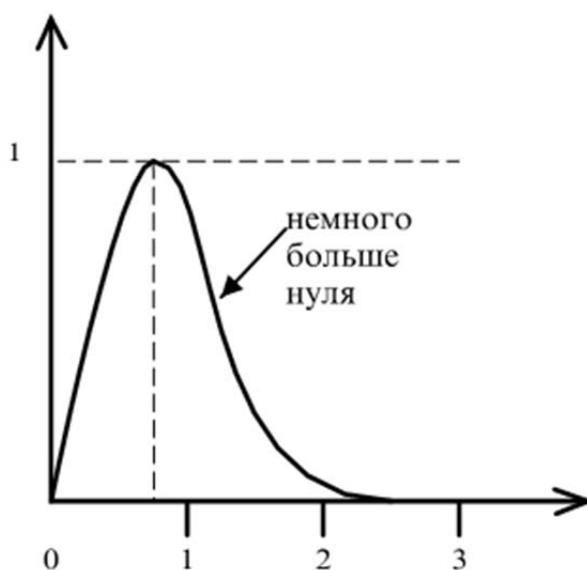
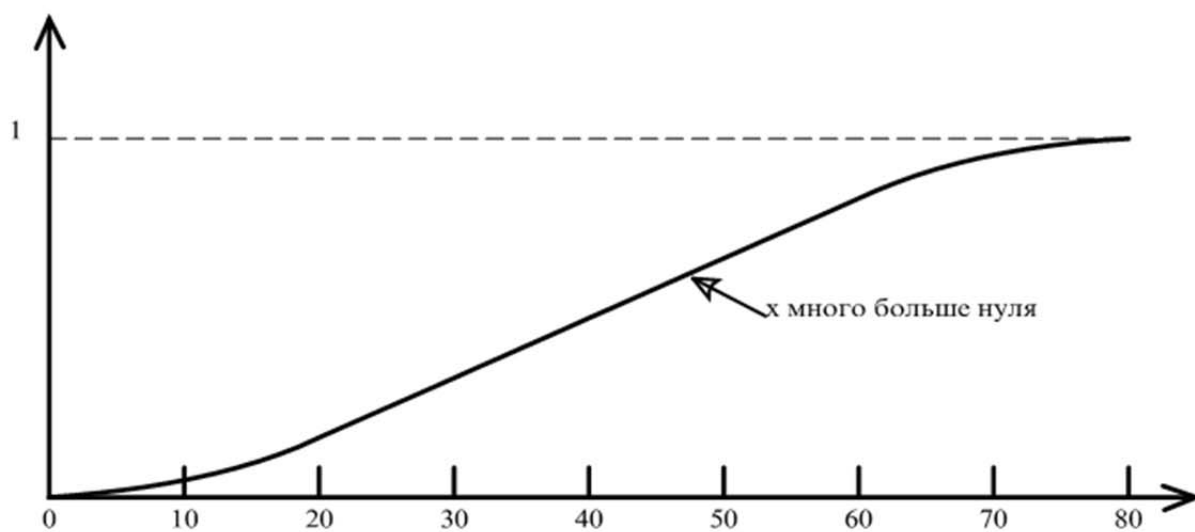
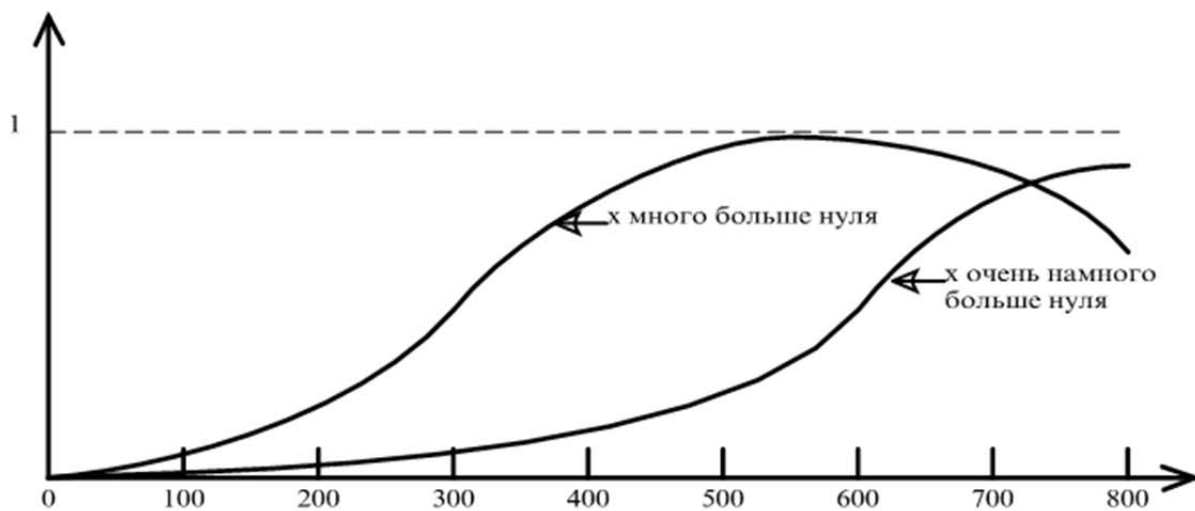


Рис. 2. Терм «немного больше a»

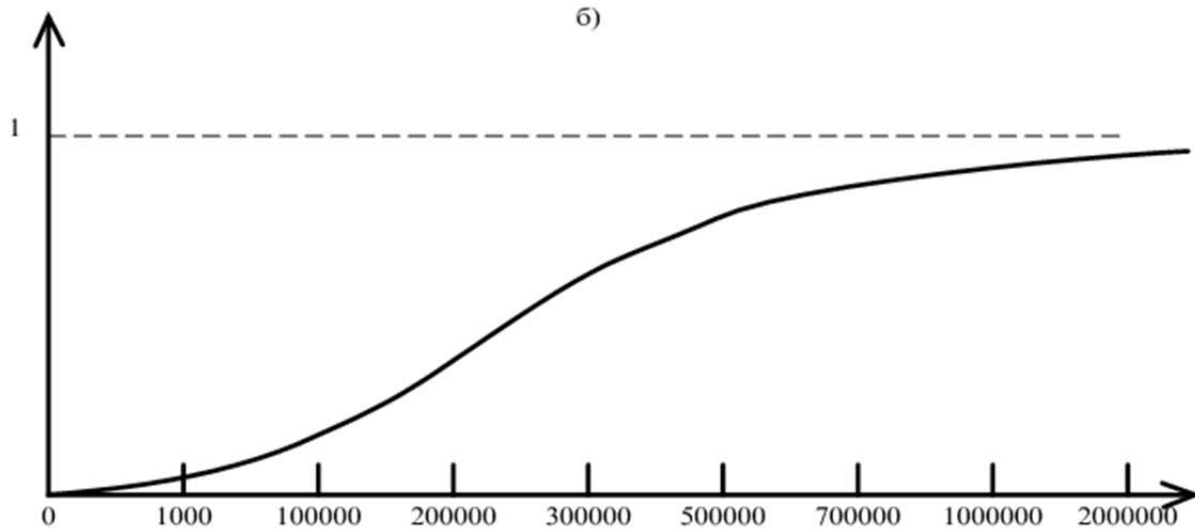
Например, понятие «X много больше нуля» имеет вид, показанный на рис. 3 а, в том случае, если проверка условия заключается в проверке принадлежности конкретного X к данному понятию и чем больше X, тем больше должно быть значение функции принадлежности к этому понятию. Если в том же алгоритме есть понятие «X очень намного больше нуля», т.е. надо проверять на принадлежность к первому или ко второму понятию, то понятие «X много больше нуля» принимает другой вид (рис. 3 б, в).



а)



б)



в)

Рис. 3. Понятия «X много больше нуля»

7.2. Семантические шкалы для измерения значений качественных понятий

Многие понятия естественного языка, которыми оперируют системы искусственного интеллекта, представляются как семантические сети, снабженные нечеткими квантификаторами типа «большой», «маленький» и пр.

Например, понятие «дом» можно характеризовать как коробку с «крышей, окнами, дверями», где понятиям «коробка», «крыша», «окна», «двери» соответствуют свои сети – определения. Можно измерить «дом» на шкале «большой – маленький» дом, окна на шкалах «большое – маленькое» окно, «много – мало» окон. Но можно измерить дом и на шкале «дом – не дом», окно на шкале «окно – не окно». Какие это будут шкалы?

Как показывают исследования психологов, люди при принятии решений пользуются большим набором шкал, крайними значениями которых являются пары взаимоисключающих понятий [2]. Как бы взвешивают принимаемое решение на множестве весов, на одних из которых написано «хороший – плохой», на других «тяжелый – легкий» на третьих «знающий – не знающий» и пр., в зависимости от ситуации. Промежуточными значениями таких шкал являются понятия «средний», «скорее хороший, чем плохой», «не тяжелый, не легкий» и пр. В теории нечетких множеств подобные понятия называются лингвистическими термами и представляются при помощи функций принадлежности, характерной для нечетких чисел.

7.3. Процедура порождения базиса нечетких термов, связанного с лингвистической парой

Ниже описываются способы нахождения среднего терма между двумя произвольными лингвистическими термами, связанными с лингвистической парой. Семантика среднего терма заключается в представлении понятия, среднего между двумя исходными термами.

Обозначения: терм будем обозначать двумя символами, первый – T , второй – символ, соответствующий ядру данного терма, например, T_a обозначает нечеткое число с ядром (т.е. максимумом) в точке a ; $\{..\}$ обозначение множества; $|...|$ – операция взятия модуля. Приведем два способа вычисления среднего терма.

7.4. Семантические шкалы для измерения значений качественных понятий

А. Нечеткие числа представлены произвольными функциями. Пусть c – точка пересечения термов T_0 и T_1 , 0 – точка максимума T_0 , 1 – точка максимума T_1 . Нахождение промежуточных термов лингвистической пары заключается в рекурсивном получении среднего терма, между двумя соседними, начиная с базовых термов на первом шаге. Средний терм $T_c = T_0 * T_1$ можно получить по следующим формулам: ядро $T_c = c$, размытие $T_c = \{T_c(x) = 1 - T_1(x) - T_0(x)\}$

Очередные средние термы T_s , получаемые из произвольных соседних термов T_c и T_r , $c < r$, можно вычислить следующим образом: пусть m – значение T_c и T_r в точке их пересечения s : $m = T_c(s) = T_r(s)$, $c < s < r$, тогда ядро $T_s = s$, где s – точка пересечения термов T_c и T_r ; размытие $T_s = \{T_s(x) = 1 - f_1(x) - f_2(x)\}$, где

$$f_1(x) = \frac{y_1(x) * (1 - m) + m}{2 - m}, \quad y_1(x) = \begin{cases} 2 - T_c(x), & \text{при } 0 \leq x \leq c, \\ T_c(x), & \text{при } c < x \leq s. \end{cases}$$

$$f_2(x) = \frac{y_2(x) * (1 - m) + m}{2 - m}, \quad y_2(x) = \begin{cases} 2 - T_r(x), & \text{при } r \leq x \leq 1, \\ T_r(x), & \text{при } s < x \leq r. \end{cases}$$

Эти операции изображены на рис. 4.

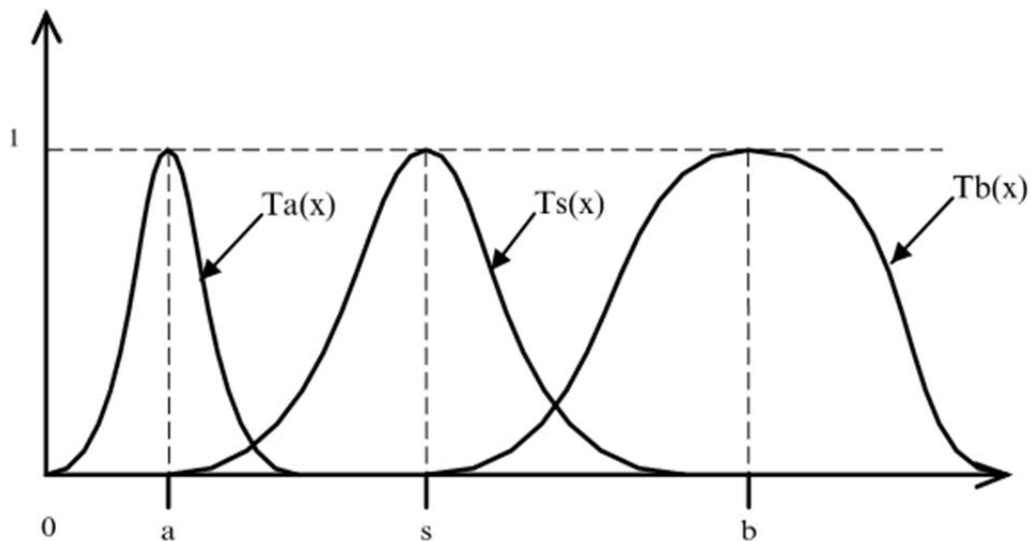


Рис. 4. Операции получения среднего терма

Б. L – R представление нечетких чисел.

Такое представление нечетких чисел было введено в работах Дюбуа – Прада [2, 3]. Левая ветвь числа определяется функцией L-типа (т.е. неубывающая), правая – функцией R-типа (т.е. невозрастающая). Тип функции – один для всего класса нечетких чисел. Левая и правая ветви результата операций над нечеткими числами – это те же L-R функции, только с другими параметрами.

Пусть $L_r(x)$ – левая ветвь $T_r(x)$, $R_r(x)$ – правая ветвь $T_r(x)$ и т.п., тогда средний терм T_s определяется функциями того же типа, L и R, соответственно, достигающими максимума в точке s , L_s и R_s .

В. Еще одна формула вычисления среднего терма для произвольного представления нечетких чисел. Очередные средние термы T_s , получаемые из произвольных соседних термов T_c и T_r , $c < r$, можно вычислить следующим образом:

Пусть m – значение T_c и T_r в точке их пересечения, s : $m = T_c(s) = T_r(s)$, $c < s < r$, тогда ядро $T_s = s$, где s – точка пересечения термов T_c и T_r ; размытие T_s вычисляется по принципу L-R представления, отдельно левая и правая часть:

$$L_s(x) = L_r(x) / m, R_s(x) = R_c(x) / m,$$

где $L_r(x)$ – левая ветвь $T_r(x)$, $R_c(x)$ – правая ветвь $T_c(x)$, / – деление обычных чисел.

Эта формула дает функцию, удовлетворяющую требованиям, предъявляемым к среднему терму пары: она достигает максимума равного 1 в точке пересечения исходных термов и обращается в 0 на границах области определения, где определены исходные термы, проще формулы А и не ограничивает вид термов, как формула Б.

Г. В любом случае средний терм T_s между двумя соседними термами T_c и T_r вычисляется на основе функционала Fu , зависящего от этих термов. Основным свойством T_s является то, что s , ядро T_s находится на пересечении термов T_c и T_r : $T_s = Fu(T_c, T_r)$, $T_c(s) = T_r(s)$.

Начав с базовых термов лингвистической пары и повторяя операцию получения среднего терма для термов, производных от этой пары, можно получить нечеткие числа, ядра которых сколь угодно близко приближаются к произвольной точке отрезка, заключенного между ядрами базовых термов. Основываясь на этом, определим для

любого значения, относящегося к той же шкале, что и базовые термы пары, способ разложения этого значения по базовым термам этой лингвистической пары.

7.5. Разложение нечеткого значения, связанного с лингвистической парой по базовым термам этой пары

Пусть e_1 – некоторое число, которое назовем порогом аппроксимации ядра, e_2 – порог аппроксимации размытия, H_k – оператор увеличения – уменьшения размытия, с параметром k , меняя который можно регулировать степень сжатия или растяжения термов.

Начав с базовых термов T_0 и T_1 при помощи описанной выше операции будем получать промежуточные термы, ядра которых расположены на минимальном расстоянии от a так, что $i < a < j$. Тогда через конечное число шагов мы получим функцию $T_b = Fu(T_i, T_j) = F(T_0, T_1)$, ядро которой примерно совпадает с ядром T_a : $|a - b| \leq e_1$; здесь F – результат неоднократного применения функционала Fu к базовым термам T_0, T_1 : ведь, начиная с первого среднего между T_0 и T_1 терма, все получающиеся после применения Fu термы в конечном итоге зависят от T_0 и T_1 . Иными словами F – это формула участия T_0 и T_1 в T_b .

Например, T_c – средний между T_0 и T_1 , тогда $T_c = Fu(T_0, T_1) = T_0 * T_1$; T_q – средний между T_0 и T_c , тогда $T_q = Fu(T_0, T_c) = T_0 * (T_0 * T_1)$; T_w – средний между T_0 и T_q , тогда $T_w = Fu(T_0, T_q) = T_0 * (T_0 * (T_0 * T_1))$;

Варьируя параметр k в операторе H_k можно изменить размытие T_b так, чтобы приблизиться к размытию T_a :

$$S(T_a) - S(H_k(T_b)) \leq e_2,$$

где S – площадь, ограниченная соответствующей кривой.

Это соотношение не всегда достижимо при произвольно заданных $H_k, T_1, T_0, e_1, e_2, Fu$. Если зафиксировано e_2 , то невыполнение этого соотношения может служить поводом для изменения H_k, e_1, Fu или T_1, T_0 . В противном случае минимальное e_2 из возможных для данного T_a будет точностью аппроксимации T_a в системе $\{T_0, T_1, Fu, e_1, H_k\}$. Таким образом, функция T_a сводится к суперпозиции базовых функций:

$$T_a \cong H_k(T_b) = H_k(F(T_0, T_1))$$

Относительное количество знаков T_0 и T_1 в функции F можно считать степенью участия базовых понятий T_0 и T_1 в формировании значения T_a . При фиксированном функционале F формула F зависит от вида термов T_0 и T_1 . Чем более размыт какой-нибудь из этих термов, тем больше относительное количество соответствующих знаков в формуле разложения F .

Еще раз заметим, что здесь рассматривается не аппроксимация нечетких множеств при помощи набора нечетких переменных, а процедура разложения нечеткого значения, связанного с лингвистической парой, по базовым термам этой пары. Процедура разложения нечетких значений по базовым термам является инструментом реализации нечетких алгоритмов. Эта процедура в корне отличается от лингвистической аппроксимации. Цель ее введения – получение для системы реализации алгоритмов «внутреннего» представления, основанного на внутрисистемных функциях T_0 и T_1 для внешней по отношению к этой системе величины – лингвистического значения T_a .

7.6. Аппроксимация логических значений

Из всего множества функций истинности выберем два представителя: $V_{\text{да}}$ и $V_{\text{нет}}$, отличающиеся от остальных функций тем, что они достигают максимумов в точках «1» и «0», соответственно. Назовем их базовыми термами «да», «нет». При помощи этих термов и определенной выше операции получения промежуточных термов можно аппроксимировать произвольную функцию истинности V : $V = F(V_{\text{да}}, V_{\text{нет}})$.

С формулой разложения лингвистического значения по базовым термам можно связать выбор названия для этого значения: название выбирается исходя из формулы разложения. При изменении базовых термов меняется формула разложения и, соответственно, название одного и того же нечеткого значения.

В нечетких алгоритмах управление режимом выполнения алгоритма может осуществляться не только при помощи логических значений, но и непосредственно самими нечеткими переменными – лингвистическими парами. Реакция на какое-то действие, т.е. нечеткое значение представленное формулой разложения на своей шкале, может быть представлена

по той же формуле, только на другой шкале – шкале реакции: формула та же, но базовые термы, участвующие в ней, с другой шкалы.

Использование лингвистических шкал обеспечивает гибкое изменение значений переменных, упрощает качественный анализ работоспособности, создает своеобразную «базу знаний». Изменяя базовые функции лингвистических шкал мы можем при помощи одного и того же алгоритма обеспечить различную величину изменений r_j .

7.7. Измерение качественных величин

В настоящее время делаются многочисленные попытки вывести некоторую универсальную шкалу измерения произвольных качеств [1–7]. Под измерением некоторой величины понимается установление изоморфизма между нею и градациями некоторой шкалы. Для измерения качественных величин используются различные шкалы: отношений, порядковые, интервалов и др. Ведется интенсивный поиск универсального представления для различных шкал измерения качества. Понятие шкалы обобщено до понятия реляционной системы. Доказано, что произвольная реляционная система изоморфна системе подмножеств упорядоченного множества, связанных отношениями ценного мажорирования. Другими словами, для произвольной шкалы измерения качеств существует универсальная форма представления, при которой пункты шкалы являются подмножествами упорядоченного множества, а структура шкалы определяется совокупностью отношений цепного мажорирования. т.е. для измерения качественной величины необходимо применить (придумать) шкалу, «отношения которой моделируют реальные отношения между объектами, а градации адекватны измеряемому качеству».

Пункты шкалы измерения качественной величины должны быть расположены так, чтобы убывание качества шло плавно, причем межклассовый переход характеризуется убыванием одного качества и нарастанием другого – его противоположности, например, «часто» переходит в «редко». Если под градациями шкалы мы будем понимать «засечки» наблюдателя, то шкала должна быть градуирована так, чтобы по мере нарастания качества количество градаций увеличивалось. Сопоставим пунктам шкалы последовательности тестовых испытаний (показаний датчиков, временных интервалов между показаниями датчиков и пр.). При этом под тестовыми испытаниями будем

понимать испытания не одного объекта, а некоторой совокупности объектов. Например, одно и то же показание датчика может входить в разные последовательности (совокупности) и вследствие этого может быть отнесено к разным пунктам шкалы.

Для измерения качественной величины необходимы:

1. проградуированная шкала измерения этого качества, подобная шкале «часто – редко» (лингвистическая шкала);

2. процедура соотнесения различных последовательностей объектов пунктам шкалы, причем под «объектами» здесь понимаются результаты как исходных испытаний (показаний датчиков, таймеров, пр.), так и результаты измерений по другим качественным шкалам. Измерение качественной величины можно считать соответствующим реальности (правильным), если воздействие на работающую систему, в основе которого лежит данное измерение приводит к желаемым результатам. Недостижение таких результатов является признаком необходимости изменения градации лингвистических шкал и процедур сопоставления последовательностей объектов пунктам шкалы.

В [1] доказана теорема о том, что измерение любого качества можно свести к измерению некоторых частных (маргинальных) качеств на порядковых шкалах. Важно, чтобы градации измеряемого качества были изоморфны градациям шкалы измерения. Многие качества, как сложные, так и простые, можно выразить переменной, колеблющейся между двумя крайними значениями, например: «сильный – слабый», «надежный – ненадежный», «годный – не годный», «часто встречается – редко встречается», и т.д. В реальной ситуации измерения абсолютные значения, которым приписываются такие лингвистические оценки, изменяются и зависят от самой ситуации измерения. Крайние лингвистические значения и все их промежуточные градации являются величинами относительными, задают только взаимный порядок их расположения, а не абсолютные (реальные) значения. Процесс сопоставления реальных значений порядковым величинам называется арифметизацией порядковой шкалы. Ниже предлагается процедура, которая одновременно и градуирует шкалу и сопоставляет ей некоторую арифметизацию.

7.8. Градуирование лингвистической шкалы

Для градуирования шкалы используем процедуру получения среднего терма из двух соседних. Пусть некоторый параметр X изменяется на

отрезке $[a, b]$. Зададим две «базовые» функции: $f_0(x)$ и $f_1(x)$: $[a, b] \rightarrow [0, 1]$. $f_0(a) = 1$, $f_0(b) = 0$, $f_1(a) = 0$, $f_1(b) = 1$, $f_0(x)$ – невозрастающая, $f_1(x)$ – неубывающая функции. Точке a припишем оценку 0, b – оценку 1, точке пересечения $f_0(x)$ и $f_1(x)$, – оценку $1/2$. Пусть f_3 – терм, средний между f_0 и f_1 , т.е. $\max f_3(x) = 1$ достигается в точке x_0 ; x_1 – точка пересечения f_0 и f_3 , тогда x_1 получает оценку $1/4$, а x_2 , точка пересечения f_3 и f_1 – оценку $3/4$ и т.д. Если X_k – точка пересечения f_i и f_j , то X_k получает оценку $A_k = (A_i + A_j)/2$, где A_i и A_j – оценки точек X_i и X_j , в которых достигают максимума функции f_i и f_j , соответственно. Точки X_k , $k = 1, \dots, N$, такие, что разница между их оценками, $A_k, A_{k+1} - A_k = \text{const} = 1/N$ будем называть градациями лингвистической шкалы A , заданной над $[a, b]$. Меняя функции f_0 и f_1 можно менять распределение градаций A .

Если рассмотрим движение точки по этим градациям то получим, что в некоторых областях точка «замедляется», в других – ускоряется. Области «замедления» естественно отнести к различным классам. Точки из «областей» ускорения можно отнести либо к двум соседним классам с малой (минимальной) степенью истинности (если они расположены между классами), либо к одному классу (крайнему) с максимальной степенью истинности (если они расположены на границах отрезка $[a, b]$) (рис. 5). На лингвистической шкале области замедления движения отслеживающей точки по оси X характеризуются увеличением плотности расположения градаций.

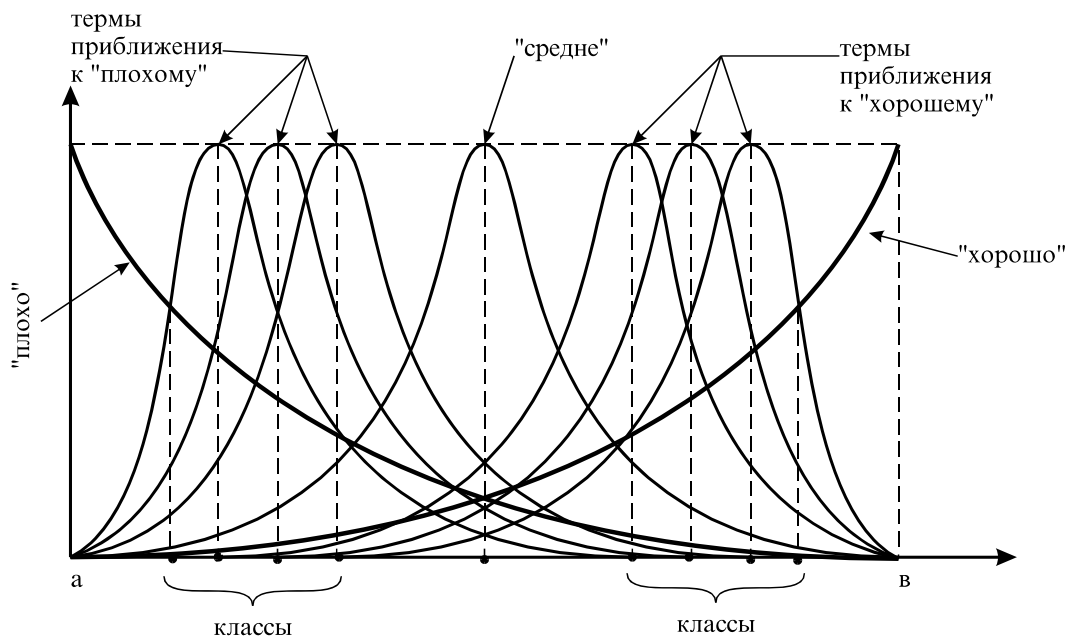


Рис. 5. Градуирование лингвистической шкалы

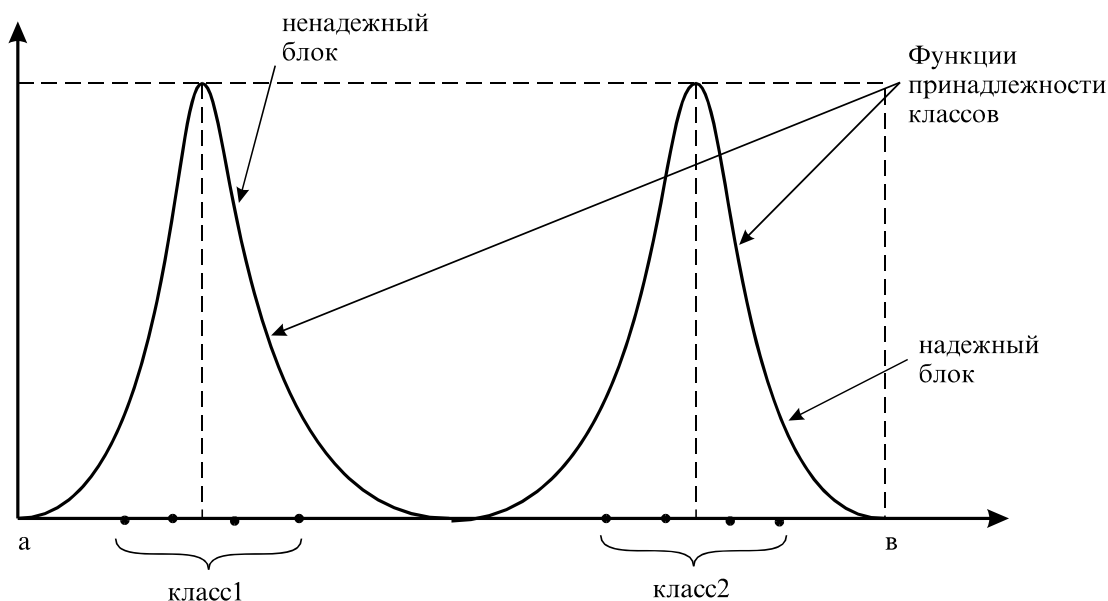


Рис. 6. Определяющие функции принадлежности к классам надежности – ненадежности по градациям шкалы

7.9. Арифметические операции над нечеткими числами

Арифметические операции над нечеткими числами определяются обычно на основе принципа обобщения Л. Заде [4]. Известно несколько определений арифметических операций, основанных на этом принципе. Разнообразие формул связано с различными ограничениями на декартово произведение в случае взаимодействия переменных. Вводятся так называемые «треугольные нормы» для того, чтобы ограничить величину размытия результатов, получаемых в соответствии с принципом обобщения Заде.

Ниже приводится определение арифметических операций, основанное на представлении о том, что размытие результата арифметической операции должно быть не меньше размытия аргументов и предположении, что размытие чисел, «почти» совпадающих с «а», должно увеличиваться при $a \rightarrow \infty$ (Например, «почти 1» это где-то от 0 до 2, а «почти 1000» – находится где-то от 900 до 1100).

Формулы, связанные с тремя треугольными нормами из четырех, рассмотренных в [3], можно получить как частный случай из этого определения, если размытие аргументов удовлетворяют условиям L-R представления. Очень вероятно, что этот результат можно было вывести на основе принципа обобщения Л. Заде, предложив соответству-

ющую треугольную норму. Однако поиск такой нормы составляет отдельную задачу, которая пока не рассматривалась.

Ниже приводится один из возможных подходов к определению операций над нечеткими числами:

Определение 7. Результат арифметической операции над простыми нечеткими числами A и B является простым нечетким числом $C = A \times B$ таким, что:

$$\text{ядро}(C) = \{X \times Y \mid J_a = \{X\}, J_b = \{Y\}\},$$

где $J_a = \{X\}$ – точечное ядро A , $J_b = \{Y\}$ – точечное ядро B .

$$\text{размытие}(C) = M_c(t) = F_a(M_a, M_b),$$

где F_a – некоторый функционал, определяющий $M_c(t)$ через $M_a(t)$ и $M_b(t)$.

Определение 8. Функционал F_a , вычисляющий размытие C , $C = A \times B$, можно задать следующим образом:

$$M_c(t) = F_a(M_a, M_b) = (f(J_c, M_a, M_b))^P,$$

где $f(J_c, M_a, M_b) ::= \max \{M_a(t - J_c + J_a), M_b(t - J_c + J_b)\}$, J_c, J_a, J_b – ядра чисел C, A, B соответственно;

$^{\wedge}$ – знак возведения в степень;

P – показатель степени, задающий степень размытия результата в зависимости от того, в какую область оси R_1 попадает результат операции. P вычисляется по следующей формуле:

$$P = \begin{cases} 1/(P_c - \max(P_a, P_b) + 1), & \text{если } x = *(\text{умножение}); \\ \max(P_a, P_b) - P_c + 1, & \text{если } x = /(\text{деление}); \\ 1, & \text{если } x = +(\text{сложение}). \end{cases}$$

Здесь P_a, P_b, P_c вычисляются по следующей формуле:

$$P = \begin{cases} \log m a_i, & a_i > m; \\ 1, & a_i < m; \end{cases}$$

где 1 может быть одним из знаков C, A, B;

m – некоторое число, определяемое экспериментально и задающее порядок изменения размытия в зависимости от ядра.

Вычитание сводится к сложению $C = A - B = A + (-B)$.

Определение 9. Одноместная операция "-" определяется следующим образом. Пусть $C = -B$, тогда:

$$\text{ядро}(C) = -\text{ядро}(B);$$

$$\text{размытие}(C) = -M_B(t) ::= M_B(-t).$$

Как показывает анализ, приведенные определения дают результаты более естественные, чем в случае обычного поэлементного определения арифметических операций над нечеткими множествами [4]:

$$M_C(X) = \bigvee_{x_1 * x_2 = X} (M_A(x_1) \wedge M_B(x_2)) \quad (1)$$

где $C = A \times B$, $\bigvee$ – операция max; $\wedge$ – операция min;

На рис. 7 представлены исходные множества A и B, а также множество $C = A \times B$ по формуле (1).

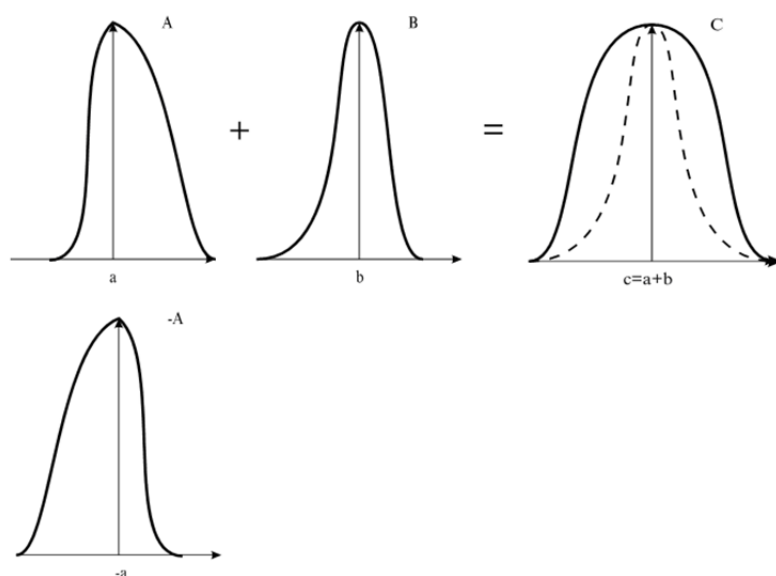


Рис. 7. Исходные множества A и B

У Л. Заде и во многих других работах [3, 4] арифметические операции сводятся к операциям над элементами (каждого с каждым) с последующей максимизацией или минимизацией значений функции принадлежности на совпадающих элементах. Такое определение имеет ограничения, особенно если иметь в виду, что функции принадлежности, определяющие нечеткие числа, имеют особый вид, который определяется путем опроса экспертов. Вид функции – результата арифметического действия, поэтому может быть весьма произвольным.

7.10. Логические переменные и операции

Логическими значениями называются простые нечеткие числа над $[0, 1]$. Эти значения (будем их обозначать $V(t)$) можно рассматривать как градации понятий лингвистической пары «да – нет», определяемой базовыми термами T_0 и T_1 .

Обоснованием выбора функций «да» и «нет» в работах Болдуина, Ягера, а также Дюбуа и Прада [3, 8] является либо процедура модификации значения при помощи функции истинности, либо схема логического вывода. А именно, сначала определяется соответствующая процедура, затем требуется, чтобы результаты, получаемые с ее помощью, удовлетворяли некоторым критериям, и на основании этого делается вывод о том, каким должен быть вид логических значений.

В настоящем разделе предлагается подход, отличающийся от изложенного. Прежде всего, определяется структура множества логических значений, т.е. выделяется базис – два специальных терма ДА и НЕТ и определяется процедура получения произвольных значений из базовых термов. Число шагов, необходимых для получения конкретного логического значения из двух базовых или, что то же, для разложения этого значения по базовым термам, зависит от вида этих термов. Подбирая ДА и НЕТ соответствующим образом, можно сократить количество шагов и/или увеличить точность представления конкретного терма. Если при вычислении предикатов получаются логические значения приблизительно из одного класса, т.е. вид их функций принадлежности примерно совпадает между собой, то за счет подбора ДА и НЕТ можно увеличить скорость и точность процедуры разложения для всех, практически вычисляемых, логических значений. Таким образом, для «стационарной» проблемной среды процедура разложения поддается оптимизации. При изменении этой среды, т.е., например, при изменении вида

числовых множеств, соответствующих тем же терминам, могут изменяться и логические значения, а значит, потребуется новая «подгонка» ДА и НЕТ для оптимизации процедуры. Следовательно, в нашем случае, обоснованием выбора ДА и НЕТ будет форма логических значений, т.е. величина, зависящая от среды реализации алгоритма, а не только от процедуры, определяемой в работе. Именно поэтому преждевременно на данном этапе исследований ограничивать вид базовых термов.

Кроме того, необходимо подчеркнуть, что в работах, посвященных принятию решений или логическому выводу функции ДА и НЕТ влияют на вывод, т.е. на вычисление предикатов. Термы ДА и НЕТ, используемые в настоящей работе, никакого отношения к этому не имеют. Значения предикатов вычисляются независимо от них. Эти понятия принадлежат к уровню интерпретации (или реализации) алгоритма. Они используются для получения внутреннего «концептуального» представления логических значений, необходимого для интерпретации нечеткого алгоритма и выражаемого формулой разложения. Подобной трактовки ДА и НЕТ не встречается в упомянутых работах.

Логические значения являются такими же функциями, как значения любых других лингвистических переменных, только с областью определения, ограниченной отрезком $[0, 1]$. Подобно тому, как функции «большой» и «маленький» отражают опыт эксперта в определенной области и могут, вообще говоря, быть и неодинаковыми, т.е. расположенными несимметрично на некотором отрезке $[a, b]$, отражающем гипотетические размеры этой области с точки зрения другого эксперта, с другой шкалой; так и функции ДА и НЕТ необязательно должны быть симметричными. В частности; воспользовавшись операцией отрицания по Заде ($T_a(x) = 1 - T_a(x)$) получим несимметричные функции ДА и НЕТ.

В нечетких алгоритмах управление порядком действий может осуществляться не только при помощи логических значений, но и непосредственно самими нечеткими переменными, связанными с лингвистическими парами.

7.11. Аппроксимация логических значений

Из всего множества функций истинности выберем два представителя: $V_{да}$ и $V_{нет}$, отличающиеся от остальных функций тем, что они достигают максимумов в точках «1» и «0», соответственно. Назовем

их базовыми термами ДА, НЕТ. При помощи этих термов и специальной операции получения промежуточных термов можно аппроксимировать произвольную функцию истинности V , как было показано выше. Ядро V будем обозначать A_v . Функция V сводится к суперпозиции базовых функций: $V = F(\text{ДА}, \text{НЕТ})$. Относительное количество знаков ДА и НЕТ в формуле F можно считать степенью участия базовых понятий ДА и НЕТ в формировании значения V . При фиксированном F_v (см. выше) функция зависит от вида термов ДА и НЕТ. Чем более размыт какой-нибудь из этих термов, тем больше относительное количество соответствующих знаков в разложении F .

Заметим, что процедура получения произвольного логического значения из двух базовых названа аппроксимацией в обычном математическом смысле этого слова, из-за того, что для ограничения количества шагов в этой процедуре приходится прибегать к приближенным соотношениям. Эту процедуру можно назвать «разложением», т.к. в результате получается формула зависимости произвольного значения от двух базовых. Процедура разложения нечетких логических значений по базовым термам является инструментом реализации нечетких алгоритмов. Эта процедура в корне отличается от лингвистической аппроксимации. Цель ее введения – это получение для системы реализации алгоритмов «внутреннего» представления, основанного на внутрисистемных функциях ДА и НЕТ для внешней по отношению к этой системе, величины – значения истинности предиката.

7.12. Логическая шкала как средство связи различных шкал между собой

Операция получения промежуточных термов определена для всех лингвистических шкал. Можно выделить шкалу для логической пары «ДА НЕТ» как основную, относительно которой задавать операции оценки для других шкал. Для этого базовые термы и определяющий функционал F_v для других шкал можно задавать в зависимости от термов и функционала для логической шкалы. Эту зависимость можно установить, например, путем отображения интервала $[0,1]$ на интервал определения другой шкалы. Таким образом, получаем, что через отображение на логическую шкалу можно установить связь между различными шкалами. Это позволит вычислять смысл таких, например, выра-

жений: «предмет такой же яркий, как другой предмет светлый» или оценивать взаимосвязь оценок разными экспертами одного и того же понятия. Таким образом, можно трактовать логическую шкалу как средство не только для вычисления степени истинности, но и для связи различных шкал между собой. Суть нечеткости – в непрерывности. Отражение на универсальную шкалу можно трактовать как своего рода логические функции, необходимые для взаимосвязи, обобщения, взаимной оценки различных понятий и смысл этого отражения в передаче непрерывности изменений от одной шкалы к другой.

Логические операции и операции сравнения нечетких множеств, определяются, например, в [2]. Способ задания этих операций зависит от интерпретации и среды реализации нечеткого алгоритма. Операции над произвольным логическим значением можно свести к операциям над ядром и размытием, как это показано выше для арифметических операций.

7.13. Подход к определению операций над нечеткими значениями, задаваемыми функциями с несколькими максимумами

Для таких функций имеют смысл следующие операции:

1. Разложение исходной функции $M_y(x)$, представляющей нечеткое понятие, на функции – нечеткие числа: выделяем точки a_i локальных максимумов, размытия чисел $T_{ai}(x)$ определяется так, что функции $T_{ai}(t)$ в точности совпадают с функцией $M_y(t)$ на участках до пересечения с соседними функциями $T_{ai-1}(t)$ и $T_{ai+1}(t)$. Набор $\{T_i(t) = V_{Ti}(t)\}$ назовем разложением $M_y(t)$.

2. Аппроксимация: функции $T_i(t)$ приближаются «лингвистическими» функциями Z , каждой из которых соответствует некоторый термин T , т.е. вместо T_i используется его приближение – результат разложения по базовым термам соответствующей лингвистической пары. В общем случае $M_y(t)$ будет представлять объединение нескольких термов Z_{ij} , соответствующих различным терминам T_j , подвергнутых операциям увеличения-уменьшения размытия и повышения-понижения достоверности.

Произвольное нечеткое множество, определяемое функцией с несколькими локальными максимумами, можно свести к лингвистическим термам и представить на некотором языке описания данных, ос-

новными элементами которого будут нечеткие числа, в свою очередь сводящиеся к небольшому набору базовых термов, связанных с некоторыми основными понятийными шкалами.

Для определения операций над нечеткими понятиями можно, если это более удобно, использовать обычное функционально-множественное представление, т.е. саму исходную функцию. Операции производятся как над некоторыми цельными объектами, используя обычные методы дискретизации функций. В случае необходимости описания результатов применяется операция разложения и лингвистической аппроксимации, описанные в пунктах 1 и 2, к множеству-результату. С другой стороны, если это удобнее, при определении операций можно учитывать возможность разделения нечетких множеств общего вида на составные элементы-числа и операции над произвольными множествами сводить к операциям над нечеткими числами. Эти операции в свою очередь, можно сводить к операциям над лингвистическими аппроксимациями чисел, т.е., в конечном счете, использовать только ядро наиболее достоверного «лингвистического числа» из объединенных при аппроксимации, а с размытием поступать как с единым образованием, не рассматривая, какие части какому термину соответствуют. Таким образом, можно определить любые операции для любых шкал. Для числовой шкалы R_1 также можно определить арифметические операции, а значит и любые методы вычислений. Для логической шкалы (шкалы истинности или уверенности) «да – нет» – логические (истинностные) значения, логические операции, операции сравнения нечетких множеств, множеств в виде объединения нечетких чисел, которые в свою очередь сводятся к базовым термам при помощи специальных операций.

7.14. Модификация значений, вычисляемых нечетким алгоритмом

Будем рассматривать нечеткий алгоритм после сборки [9], когда этот алгоритм представляет из себя последовательность операторов и условий. Будем считать также, что переменные, используемые в алгоритме, принимают только такие значения, которые сводятся к определенным выше нечетким числам. Выполнение операторов осуществляется при помощи системы интерпретации, которая определяет, в ка-

ком виде выполнять данный оператор и как изменить результат его выполнения. Для модификации операторов и значений в общем случае может быть использована автономная база знаний, способная меняться как независимо от выполняемого алгоритма, так и благодаря ему. Ниже мы опишем некоторый простой метод «функциональной» модификации, когда значения-функции, вычисляемые операторами изменяются в соответствии со «степенью уверенности», определяемой логическими значениями V , вычисленными на предыдущих шагах выполнения алгоритма.

Будем считать, что степень уверенности в том, что необходимо делать действия, указанные на ветке, соответствующей логическому значению V , равна степени совпадения V с понятием «да», степень неуверенности – совпадению с «нет». Предположим, что «уверенность» будет сказываться в «уточнении» результатов выполнения операторов (т.е. нечеткое множество будет «сжиматься» возле точек локальных максимумов), «неуверенность» – в понижении «достоверности» результатов (максимумы нечеткого множества будут пропорционально понижаться). Выше было показано, как любую логическую функцию можно представить в виде некоторой комбинации базовых термов ДА и НЕТ. Поэтому операцию модификации при помощи произвольной функции истинности можно свести к операции модификации при помощи базовых термов ДА и НЕТ.

Определение 10. Модификацию произвольного нечеткого числа T базовыми логическими термами определим следующим образом:

$$T(x)|_{da} = H_p(T(x)),$$

$$T(x)|_{net} = (T(x)/m) \wedge 1,$$

где H_k – оператор увеличения – уменьшения размытия, p – некоторое фиксированное значение параметра k , при котором оператор H вызывает уточнение $M(x)$, $M > 1$ – некоторое число, индекс уменьшения, соответствующий НЕТ.

Определение 11. Пусть $V = F(ДА, НЕТ)$ – произвольная функция, получаемая из базовых термов ДА и НЕТ при помощи операции получения среднего терма, определенной выше. Тогда модификацию при помощи V можно определить следующим образом:

$$T(x)|_v = ((T(x) \wedge (p \wedge i)) / (m \wedge j)) \wedge 1,$$

где i, j – относительное количество знаков ДА и НЕТ, соответственно, в функции F . Например, если

$$V = F(\text{ДА}, \text{НЕТ}) = ((((((\text{ДА} * \text{НЕТ}) * \text{ДА}) * \text{ДА}) * \text{ДА}) * \text{ДА}) * \text{ДА}), \text{то}$$

$$i = 6/7, j = 1/7, T(k)|_v = ((T(x) \wedge (p \wedge (6/7))) / (m \wedge (1/7))).$$

Модификация при помощи произвольного логического значения $V_v = H_k(V)$ будет выглядеть как модификация при помощи соответствующего V^* с поправкой k , а именно:

$$T(x)|_{v_v} = T(x)|_{H_k(V)} = ((T(x) \wedge (k * p \wedge i)) / ((m \wedge j) / k)) \wedge 1.$$

Выбор базовых термов ДА и НЕТ определяет вид разложения F , это отражается на степени модификации результатов выполнения операторов $T(x)$. Значит, изменяя вид ДА и НЕТ можно изменить интерпретацию алгоритма. Например, можно предложить следующий подход. Пусть результат проверки каждого условия сказывается на степени уверенности автомата в правильности выбранного пути движения по алгоритму. Если интерпретатор «уверен» в том, что нужно выполнять операторы, то одну и ту же функцию V он будет склонен принимать за «да», чем в том случае, когда он «неуверен». Каждый ответ «да» на пути движения автомата повышает его «уверенность», каждый ответ «нет» повышает его «неуверенность». Реализовать это можно при помощи модификации базовых термов $V_{\text{да}}$ и $V_{\text{нет}}$. Повышение «уверенности» будет заключаться в расширении терма $V_{\text{да}}$, повышение «неуверенности» – в расширении $V_{\text{нет}}$. Чем больше «уверенность» автомата, тем в большей степени будет уточняться $T(x)$; чем больше «неуверенность», тем в большей степени будет понижаться достоверность $T(x)$ при одном и том же значении V .

Модификацию базовых термов ДА и НЕТ произвольным логическим значением V так же можно свести к модификации базовыми термами.

Можно предложить другой тип модификации $T|_v$ в основе которого лежит другая трактовка изменения вида нечетких значений в зависимости от степени уверенности автомата.

Определение 12. Нечеткое значение определим как «нечеткий коридор» ширины h вокруг функции – терма, определенного выше. Уточнение и размывание этого значения выражается в уменьшении или в увеличении, соответственно, ширины коридора.

Величина h может быть характеристикой степени «доверия к этому терму». В операциях сравнения считается, что это значение совпадает с термом, если терм попадает в его «коридор».

Тогда модификация T будет заключаться в изменении только одного числа h в зависимости от формулы разложения V по базовым термам пары ДА-НЕТ, она определяет «доверительный интервал», две функции T_1 и T_2 – нижнюю и верхнюю границу возможных значений.

Изложенный подход можно расширить для определения модификации произвольных термов при помощи значений любой лингвистической пары, а не только логических значений.

Разделение чисел на ядро и размытие дает возможность определить эти операции как объединение операций двух типов: над ядрами и над размытиями, что сказывается в упрощении определения операций и может дать существенную экономию оборудования при построении цифровых систем обработки семантической информации. Нечеткие логические значения оказываются подмножеством нечетких чисел. Это позволяет унифицировать представление нечетких величин, анализировать связи между нечеткими понятиями и подойти к определению процедур, соответствующих концептуальным зависимостям естественного языка. Изложен процесс модификации нечетких понятий при помощи логических значений (понятий шкалы истинности), моделирующей влияние степени уверенности на результат выполнения действий, допускающий расширение для определения модификации при помощи значений из других шкал. Модификация текущих значений может использоваться вычислительной системой, выполняющей алгоритм, для настройки исполнительных механизмов на текущую ситуацию, а изменение базовых термов представляет собой более глубокий уровень адаптации в системе. Рассмотренный метод модификации позволяет наметить один из возможных подходов к построению адаптивных самоорганизующихся систем реализации нечетких алгоритмов и обработки семантической качественной информации.

Выше была описана модификация результатов выполнения оператора на основе нечеткой информации об истинности предиката, связанного с этим оператором. Подобная модификация отражает такое свой-

ство деятельности человека, как влияние уверенности в необходимости (или нужности), некоторого действия на результаты этого действия. Например, если человек уверен, что его работа «никому не нужна», то он и делать ее будет «спустя рукава». Предлагается указанную степень уверенности (и неуверенности) вычислять в зависимости от истинности предиката и вида базовых термов пары «да – нет».

7.15. Анализ работоспособности резервированных систем на основе качественной информации, представленной нечеткими числами

Высокий уровень сложности современных технических систем приводит к тому, что традиционные точные методы анализа работоспособности требуют усовершенствования или дополнения некоторыми качественными методами, которые ценой потери точности позволяют решить задачу обеспечения надежности в реальный промежуток времени в целом. Управляющая система в этом случае должна быть снабжена некоторыми элементами искусственного интеллекта, обеспечивающими возможность работы с качественной информацией, и должна обладать адаптивными возможностями для уточнения этой информации при проверке соответствия между своими «внутренними представлениями» и «реальными событиями».

Моделирование функций мозга человека является одним из основных приемов, используемых при создании систем с искусственным интеллектом. Представление о надежности у человека, наблюдающего за поведением системы изменяется, по-видимому, следующим образом: вначале, когда он запускает новую, отлаженную систему, он уверен, что она надежна. После серии сбоев уверенность в надежности системы начинает медленно падать. Затем наступает некоторое «насыщение», после которого даже единичный или небольшое, сравнительно, количество сбоев вызывает резкое изменение представления о надежности. Человек от относительной уверенности в надежности переходит как бы скачком к относительной уверенности в ненадежности системы (блока). Затем уверенность в ненадежности медленно крепнет и наконец доходит до предела за которым необходима реконфигурация. Для градуирования шкалы измерения надежности используются базовые функции и процедура градуирования, определенные выше. Отслеживающая точка, движущаяся по

оси X от надежности к ненадежности проходит области большей или меньшей плотности градаций. Там, где больше плотность градаций, движение отслеживающей точки как бы замедляется, там, где плотность градаций меньше – ускоряется. Области замедления движения этой точки можно классифицировать терминами «надежный класс (блок)» и «ненадежный блок». Функции, определяющие эти классы в соответствии с градациями лингвистической шкалы рис. 5 показаны на рис. 6.

Работоспособность блока определим как качественную величину, описывающую (формализующую) представление экспертов о работоспособности: «чем чаще сбоит блок, тем ниже его работоспособность». Для измерения этой величины необходимо формализовать, что значит «частота сбоев блока» или «уровень сбоев».

Если определить частоту сбоев на текущий момент как количество сбоев блока за какой-то промежуток времени до этого момента, то встает вопрос, каким должен быть этот промежуток, должна ли величина его меняться и по какой причине. Кроме того, на основании представления о работоспособности, человек делает выводы о пригодности блока к дальнейшей работе. Следовательно, при определении «уровень сбоев» блока в текущий момент времени важна не столько предыстория, сколько прогноз его поведения в «ближайшем будущем». Многие методы прогноза основаны на построении аналитического выражения для функции, наилучшим образом «приближающей некоторый отрезок предыстории». Будущие значения вычисляются по полученной формуле. В общем случае выбор формулы достаточно сложен и приводит к решению трансцендентных уравнений. На практике обычно употребляются формулы первого, реже второго порядка. Чтобы увеличить точность прогноза, функцию на разных участках приближают при помощи различных выражений и используют методы выявления закономерностей для выбора (вычисления) аналитического выражения на очередном участке. Подобные методы в случае больших требований к точности прогноза приводят к значительному объему вычислений и для выявления закономерностей нуждаются в просмотре предыстории на значительную глубину. Кроме того, определить «частоту сбоев» блока в данный момент – это не то же самое, что достаточно точно предугадать величину интервала между очередными сбоями. Понятие «уровень сбоев» (или понятие «работоспособность»), которым пользуется человек в своей нечеткой формулировке скорее всего обозначает некоторую интегральную характеристику, в которой учитывается и средняя величина интервала между сбоями за

некоторый период времени, и прогноз очередного интервала и предусматривается предупреждение возможности появления недопустимо малых промежутков между сбоями (т.е. случай, когда система не успевает восстановить информацию после очередного сбоя, как происходит следующий сбой – так называемую «ситуацию отказа»).

Таким образом, понятие работоспособность блока – это качественная характеристика, для вычисления которой необходима система, способная учитывать все смыслы, заложенные в это понятие, способная менять свои внутренние параметры для выделения среди этих смыслов главного в текущий момент времени (в зависимости от «состояния» блока). Такая система должна обладать свойством саморегуляции для усовершенствования своей внутренней структуры.

7.16. Структура шкалы для измерения работоспособности блока

Структуру качественной шкалы для измерения работоспособности определим следующим образом. Шкала представляет собой объединение с одной стороны «лингвистических шкал», градации которых задают степень изменения качественной величины работоспособности по мере увеличения или, соответственно уменьшения качества, а с другой стороны – процедуры сопоставления различных последовательностей временных интервалов между сбоями пунктам «лингвистической шкалы». В отличие от обыкновенной шкалы, пункты которой равноудалены друг от друга, пункты лингвистических шкал расположены неравномерно. Градации этих шкал задаются при помощи процедуры градуирования, определенной выше. Для сопоставления последовательностей временных интервалов пунктам шкалы используются процедуры прогноза очередного интервала между сбоями блока на основании нечеткой информации о предыстории, обобщающей предыдущую последовательность сбоев. Для вычисления работоспособности будем применять три лингвистических шкалы: L_t , L_{t+} , L_{t-} .

7.17. Анализ работоспособности блоков

Рассмотрим задачу анализа работоспособности блоков B_j , $j=1...q$ системы с холодным резервированием, выполняющей некоторый

набор функций f_n , $n = 1 \dots q$. Работоспособность B_j зависит от частоты сбоев в единицу времени. Информация о неисправности (сбое или отказе) появляется как ненулевой набор показаний датчиков M_i , $i = 1 \dots p$. Локализовать неисправность с точностью до блока не удастся ввиду недостаточности современных средств диагностики.

Функции, выполняемые системой, разделены на группы по важности. Одни и те же блоки могут выполнять разные функции. Режим использования блоков определяется тем, на какую функцию он будет подключен в «момент переключения». Один и тот же блок может работать хуже или лучше в зависимости от условий эксплуатации. Необходимо обеспечить работу системы на достаточно надежных блоках, причем более важные функции должны выполнять более надежные блоки в соответствии с условиями среды и режимом работы системы.

Необходимо обеспечить выполнение f_n на блоках B_i с приемлемым для каждой f_n уровнем работоспособности r_j . В процессе работы возможна реконфигурация системы, т.е. перераспределение блоков для выполнения других функций и подключение новых блоков из холодного резерва.

Опишем процесс вычисления значений r_j , $r_j \in [0, 1]$, работоспособности блоков B_j при условии, что работоспособность блоков из холодного резерва считается неизменной; M_i , $i = 1 \dots p$, считаются символами только о сбоях и r_j зависит только от S_{ij} – силы связи между M_i и B_j и от t_j – величины интервала между двумя последними сбоями, причастными к B_i , т.е. такими M_{i1} , M_{i2} , что $S_{ij} < 0$ и $S_{ij} > 0$. r_j , $j = 1 \dots q$ будем изменять при помощи лингвистических шкал L_{r+} и L_{r-} . L_{r+} – для увеличения r_j , L_{r-} – для уменьшения r_j . Значение r_j оцениваем по L_{r+} или L_{r-} , соответственно (обозначаем L_r).

Оценка r_j по L_r : Пусть r_j попадает в интервал $[a(k), a(k+1)]$ шкалы L_r . Если r_j расположен ближе к $a(k_r)$ или точно по середине между $a(k)$, $a(k+1)$, тогда r_j получает оценку $a(r_j) = a(k)$, в противном случае $a(r_j) = a(k+1)$.

Изменение r_j : Пусть $a(r_j)^+$ – оценка r_j по L_{r+} , $a(r_j)^-$ – оценка r_j по L_{r-} . Уменьшение r_j зависит от величины градации $V(a(r_j)^-)$, расположенной слева от $a(r_j)^-$ для L_{r-} , а увеличение r_j зависит от величины градации $V(a(r_j)^+)$, расположенной справа от $a(r_j)^+$ на шкале L_{r+} .

Оценка t_j : Долю от $V(a(r_j)^-)$, на которую уменьшается r_j в зависимости от t_j определяем при помощи оценки t_j на шкале L_t . Эту оценку определяем следующим образом: пусть t_j попадает в интервал $[a(kt)$,

$a(k+1)t]$ шкалы L_t , тогда t_j получает оценку $a(t_j) = a(t_k)$. При такой оценке малые значения t_j оцениваются нулем и не будут влиять на изменение r_j , а большие значения t_j будут сокращать уменьшение r_j . Значения, отражаемые на этой шкале, должны зависеть от оценки среднего интервала между сбоями, характерного для этого блока.

Пригодность блока с работоспособностью r_j для выполнения функции f_n будем определять при помощи оценок пригодности $M_{fn}(r_j)$. Значение $M_{fn}(r_j)$ должно быть больше некоторого уровня B , чтобы B_j можно было подключить для выполнения f_n . При задании областей подключения учитывается важность f_n и возможность выполнения данной функции на блоках с соответствующим уровнем сбоев (уровень сбоев оценивается значениями r_j). При реконфигурации для выполнения функции f_n выбираются блоки из тех, что остались неподключенными. Реконфигурация считается успешной, если для каждой f_n нашлись блоки B_i с таким r_j , что $M_{fn}(r_j) > B$.

Обозначим: $E_j, j = 1...q, E_j \in \{1...1\}$ – показатели подключения блоков; $E_j = n$, означает, что блок B_j подключен для выполнения функции f_n ; $E_j = 0$ (блок в резерве); M – показатель сбоя; $M = 0$ – значит сбой нет; $M = i, i = 1...q$ – значит произошел сбой M_i . Тогда алгоритм вычисления r_j будет выглядеть следующим образом:

1. Задаются начальные данные: $S_{ij}, E_j, r_j, t_j = 0, i = 1...q, j = 1...g$.
2. Увеличивается t_j : $t_j := t_j + 1, j = 1...q$.
3. Если $M = 0$, то на 12 (на увеличение r_j).
- {Обработка сбоя (уменьшение r_j)}
4. $i := M$, определяется множество J номеров блоков, причастных к M : если $S_{ij} = 0$, то j записывается в J .
5. Оценка t_j по шкале L_t : определяется $a(t_j), j \in J$.
6. Оценка r_j по шкале L_r : определяется $a(r_j)-, j \in J$.
7. Определяется $V(a(r_j)-), j \in J$.
8. Уменьшается r_j : $r_j := r_j - S_{ij}(1-L(t_j)) * V(a(r_j)-), j \in J$.
9. Проверяется соответствие блоков выполняемым функциям: $M_j := M_{fn}(r_j)$, если $E_j = n$; если $M_j > B$, для всех $j \in J$, то $M := 0, t_j := 0, j \in J$, переход на 2.
10. Реконфигурация.
11. Если реконфигурация успешна, то $M := 0$, устанавливаем новые значения $E_j, t_j := 0, j \in J$ переход на 2.
12. {Увеличение r_j }

13. Определяется множество I номеров блоков, для которых время работы превысило некоторый порог, $t_j > t_p$: для всех $j = 1 \dots q$, если $t_j > t_p$, то j записывается в I ; если I пусто то переход на 2.

14. Оценка r_j по шкале L_{r+} : определяется $a(r_j)+$, $j \in I$.

15. Определяется $V(a(r_j)+)$, $j \in I$.

16. Увеличивается r_j : $r_j := (r_j + V(a(r_j)+)) \bmod 1$, $j \in I$.

17. $t_j := t_j - t_p$, $j \in I$, переход на 2.

Использование лингвистических шкал обеспечивает гибкое изменение значений переменных, упрощает качественный анализ работоспособности, создает своеобразную «базу знаний». Изменяя базовые функции лингвистических шкал, мы можем при помощи одного и того же алгоритма обеспечить различную величину изменений r_j .

7.18. Методика опроса экспертов для выяснения значений S_{ij} – степени причастности блока B_j к сбою M_i

Экспертам задается вопрос: Какие блоки могут быть причиной сбоя M_i ? Ответ формулируйте как высказывание на привычном для вас языке. Возможные ответы: да, нет, возможно, скорее всего причастен, скорее всего не причастен, может быть и не причастен, мало причастен, возможно и такое, и т.д. Расположите их в порядке убывания причастности. Там, где вам кажется необходимым, определите близость соответствующих высказываний при помощи определений. Например: почти совпадают, очень близки, очень далеки.

В результате получаем серию примерно таких ответов: причиной сбоя M_i могут быть следующие блоки

ответ первого эксперта: блок 1: да; блок 2: обязательно; блок 3: возможно; блок 4: никогда. (1, 2 почти совпадают)

ответ второго эксперта: блок 1: обязательно; блок 2: возможно; блок 3: вероятно; блок 4: может быть; блок 5: скорее всего нет; блок 6: нет; (2, 3, 4 близки; 5, 6 почти совпадают).

7.19. Построение s_{ij}

S_{ij} – это логическое значение, $S_{ij} \in [0, 1]$.

1. Выбор шага расположения оценок S_{ij} на шкале зависит от ситуации, т.е. от количества блоков, от количества групп близости определенных экспертом. Например, для оценки второго эксперта можно принять следующее: Шаг $1/6$ – стандартный шаг; $1/12$ – шаг между «близкими»; $1/32$ – шаг между «почти совпадают».

2. Численная оценка качественных (словесных) ответов заключается в следующем:

а. разбиение словесных ответов на группы по близости в соответствии с определением экспертов;

б. оценка групп, т.е. расположение их на лингвистической шкале;

с. оценка каждого элемента, соответствующего конкретному блоку.

Например, в ответе второго эксперта выделяются 3 группы качественных понятий: 1 – обязательно; 2 – возможно, вероятно, может быть; 3 – скорее всего нет, нет.

ФОРМУЛА: группы располагаются между 0 и 1 равномерно, члены группы вокруг центра группы с шагом, соответствующим данной группе. Т.о., каждый «ответ» получает численное выражение S_{ij} .

В случае если S_{ij} оценивается на основании мнения коллектива экспертов, то имеет смысл поступить следующим образом:

1. Упорядочить $\{\text{Эк}\}$ по степени компетентности: $h_k > h_{k+1}$, где h_k – оценка компетентности k – ого эксперта;

2. За основу – мнение «наиболее компетентной кучки». Например: Эк_1 принадлежит «очень компетентной», т.е. к первой группе, остальные – к «достаточно компетентной» и т.д. следовательно: S_{ij} оценится следующим образом: {скорее всего сила связи B_i , M_j равна S_{ij1} , очень возможно, что равна S_{ijk} , k принадлежит «достаточно компетентной»; может быть равна S_{ijk} , k принадлежит «не очень компетентной»}.

Возможна ситуация, когда единого «нечеткого числа» выделить нельзя – мнения разделяются, тогда получается столько, сколько надо, за основу (за ядро) всегда берется S_{ijk} , k принадлежит «очень компетентной».

Уровень дискретизации – произвольный от 1 числа на всю группу, до индивидуального N_k для каждого числа S_{ijk} .

Два нечетких числа S_{ij1} , S_{ij2} отражающие мнение двух экспертов объединяются в зависимости от оценки компетентности экспертов приписыванием им лингвистических оценок: сила связи B_i , M_j скорее всего S_{ij1} , может быть S_{ij2} если $S_{ij1} > S_{ij2}$, где $>$ – отношение предпо-

чтения, определяется по h_k ядра. При вычислении g_i используется наиболее предпочтительный S_{ij} .

Смысл S_{ij} : а) выделяет группу экспертов, «ответственных» за решение о силе связи между B_i и M_j (опосредованно через «шкалу» h_k); б) определяет набор возможных значений «ядра», определяет предпочтительность каждого значения: $Y(S_{ij})$ – это то конкретное числовое значение, которое можно использовать вместо S_{ij} в обычных вычислениях.

7.20. Алгоритм построения s_{ij}

Исходные данные: S_{ijk} – трехмерный массив, h_k одномерный массив, k принадлежит K – множество (количество) экспертов. Задача: сгруппировать S_{ijk} так, чтобы получить одно или несколько нечетких чисел, т.е. выделить группы $S_{ijl} = \{S_{ijk}, k \in K\}$, $l \in L =$ множество (количество) групп.

1. фиксируем i, j , k меняется цикл по k (K -упорядочивание S_{ijk} : $S_{ijk1} < S_{ijk2}$ если $h_{k1} < h_{k2}$, тогда по h_k можно определить номер эксперта, « k »).

2. выделение ядер $y_l = \text{ядро}(S_{ijk}, h_k)$, $h_k = \max \{h_k \mid (S_{ijk}, h_k) \text{ не принадлежит } S_{ijl}, l = 1, 2, \dots, l-1\}$.

3. Выделение ядра и построение первой функции S_{ijl} .

$Y_1 = S_{ijk} - h_k = 1$ ($\max(h_k \mid h_k \text{ принадлежит достаточно компетентным})$).

Метрика на множестве S_{ijk} – функции «почти (x)» определенные на всем $R : (S_{ijk}, h_k)$, которые закрываются функцией «почти (y_1)» входят в группу S_{ijl} .

В общем случае представление о работоспособности, g , должно выглядеть как «нечеткое число – понятие», т.е. функция, а не просто число. Изменение g в соответствии с «качественным» методом, определенным выше – это «кирпичики» в построении g . Размытие числа g может определяться как результат менее достоверных оценок при помощи параллельно действующих процессов прогноза, как результат опроса группы экспертов или другим способом, например, так:

Ядром числа g будет координата на оси X лингвистической шкалы L_g , размытием – локальный прогноз изменения в «ближайшем будущем» определяемый изменением ядра в случае, если произойдут (правое крыло R) или не произойдут сбои (левое крыло L).

Величина размытия определяется размером градаций шкалы в окрестности r , каждая градация соответствует изменению «представления о надежности» в случае, если произойдет сбой (градации слева от r). На лингвистической шкале отмечается только ядро r . Представление r как нечеткого числа осуществляется в других координатах: на оси X – окрестность ядра r со шкалы, на оси Y – отметки «времени». Каждому типу внешних условий для каждого блока сопоставляется свое значение r .

Преимущество такого нечеткого числа r_j – в возможности учета локального прогноза в момент реконфигурации и подключения новых блоков.

Реализация r как нечеткого числа – понятия возможна на матричной структуре с «ограниченным периодом памяти»: т.е. через определенный период «значение» съезжает вниз. «Каждый столбик – как термометр» можно реализовать отдельно, как автономный вычислительный процесс.

При реконфигурации значения r_j оцениваются по другим лингвистическим шкалам – шкалам подключения, отличающимся видом базовых функций, определяющих распределение градаций на шкале. На этих шкалах выделены «области подключения», определяющие наиболее предпочтительные значения r_j при выборе блоков для выполнения соответствующих функций. Для каждой функции своя шкала. Абсолютные значения на всех шкалах одинаковые. В случае конкуренции блок подключается для выполнения более важной функции. В этом случае для выполнения менее важной функции возможно подключение менее надежного блока, чем «указано» в «области подключения» для данной функции. Если есть выбор, то подключается блок со «средним» значением r_j из «области подключения» для того, чтобы «был запас» допустимых блоков.

Предлагаемый подход обеспечивает возможность изменения параметров управляющей системы, дает возможность «подгонки» параметров, отражающих некоторые качественные представления (о надежности блоков, о важности функций, о допустимых параметрах надежности для выполнения функций и пр.) к реальному, получаемому опытным путем в процессе эксплуатации системы «положению вещей».

Тестирование, т.е. принятие решений о запуске тестов, можно реализовать аналогично с принятием решений о работоспособности

блоков. Так же как S_{ij} – коэффициенты связи блока B_i с сообщением о сбое M_j , можно организовать представление C_{ik} – коэффициентов связи блока B_i с тестом T_k . Происхождение этих коэффициентов одинаковое – отражение опыта экспертов. Вид их так же нечеткое число-понятие отражающее не только осредненное мнение, но и все мнения, в масштабе (т.к. $\max < 1$)

Связь S_{ij} можно представить как связь множества блоков B_i с множеством сообщений о сбоях M_j – получается понятие – поверхность. Сечения по каждому B_i дадут кривую – число-понятие; если усечь до точки, то получится число-коэффициент, т.е. число принадлежащее $[0, 1]$ – осредненное, достаточно грубое, без вариантов.

Метод измерения качества при помощи движения отслеживающей точки по лингвистической шкале с изменением и настройкой этих шкал в зависимости от результатов дает возможность менять число – коэффициент и работая с обычными числами выходить на уровень нечетких понятий.

Заметим, что S_{ij} , C_{ik} не отражает полноты картины. Функция-поверхность должна быть нечеткой, т.е. представлять из себя коридор с внутренней структурой такой, чтобы при усечении до точки (B_i , M_j) появлялась функция: число-понятие, т.е. для отражения нечеткости необходимо не трехмерное, а четырехмерное пространство. Таким образом, S_{ij} должно быть числом-понятием на лингвистической шкале. Понятие необходимо, потому что есть отражение «еще одной координаты» – экспертов. Те или иные результаты позволят корректировать «отношение к экспертам» – т.е. повышать коэффициент важности мнения тех, по чьим оценкам хорошо работает система. S_{ij} отражает статистику мнений экспертов о взаимосвязи блока B_i и сигнала M_j . В другой ситуации (внешние условия, др.) статистика мнений может быть другой, т.о. ситуация сказывается в изменении внешнего вида функций принадлежности. C_{ik} – тоже отражение статистики.

Для принятия решений о реконфигурации рассмотрим задачи: 1. прогноза очередного интервала между сбоями. 2. оценки среднего интервала между сбоями для подсчета, «как сбоит блок».

Цель первой – это как можно точнее угадать очередной интервал вычислить динамику поведения всего процесса, вывести формальную зависимость – это трудно и для реального процесса возможно только как определение динамического процесса смены формул прогноза в зависимости от того, какая из возможных формул оправдывается в

текущий момент. Подход к построению такой системы был приведен выше. Цель второй – Оценить (вычислить) среднее значение, взятое по окрестности текущей точки, т.е. вычислить «текущее среднее», на основе которого при помощи динамического изменения оценок по разным шкалам определять прогноз поведения системы в ближайшем будущем. Соответствие реальному процессу обеспечивается изменением базовых функций, соответствующих градаций и «текущего окна» для набора лингвистических шкал, на основе которых делается вывод и принятие решений.

7.21. Прогноз интервала между сбоями

Рассмотрим динамическую смену прогнозов, вычисляемых методом средних квадратов по формуле $y = kx + b$. Чем точнее «в среднем» мы указываем величину очередного интервала между сбоями, чем «ближе к сбою» мы можем расположить контрольную точку – тем меньше будут потери: 1) рабочего времени, теряемого при возвращении в контрольную точку; 2) времени на перекачку информации – так как чем меньше промежуток от контрольной точки до сбоя, тем меньше за это время возможно обращений в память. Если n -тая контрольная точка установлена раньше или позже того момента, когда произошел n -й сбой, то происходят потери информации. Точно установить контрольные точки невозможно, следовательно, цель заключается в том, чтобы минимизировать потери при помощи динамического прогноза. Будем «запускать» параллельно несколько процессов прогноза, каждый из которых будет вычислять прогноз по формуле, например, наименьших квадратов, но на другом промежутке. Для каждого из этих процессов будем считать S_1 – суммарное время недостатка величины контрольного промежутка до интервала между сбоями и S_2 – суммарное время превышения величины контрольного промежутка над интервалом между сбоями. Обозначим f_n величину, на которую допустима ошибка в оценке интервала между сбоями, т.е. величину «коридора» в который должна попасть контрольная точка по отношению к реальной точке сбоя.

Принимать решения об изменениях контрольной точки можно только после «приобретения уверенности, что ситуация действительно такова, как кажется», например, если «реальный интервал постоянно меньше прогноза» то следует менять прогноз, чтобы уменьшить потери.

Критерии принятия решений Перечислим возможные решения:

1. Изменить прогноз интервала между сбоями (посчитать новые k и b) или «запустить новый процесс прогноза», т.е. начать подсчет новых коэффициентов для вычисления новой формулы по которой в дальнейшем будет осуществляться прогноз. Новая формула активизируется, если «возникает уверенность, что менять надо», эти коэффициенты обнуляются, если «возникает уверенность, что менять не надо» (в смысле оценки по соответствующей лингвистической шкале).

2. Установить (изменить) «коридор для контрольной точки», т.е. f_n ;

3. Осуществить реконфигурацию, если надежность блока понизилась ниже такого уровня, когда на перекачку в новые блоки понадобится столько времени, что даже если будем работать с меньшим количеством сбоев в единицу времени, то не уложимся в отпущенное для решения задачи время.

4. Тестировать. Вопрос о запуске тестов так же можно решить, пользуясь нечеткими оценками.

Возможные критерии принятия решений: 1. минимум допустимых потерь информации при сбое. 2. минимум потерь на восстановление при сбое. 3. минимум отказов. 4. минимум сбоев. 5. выполнение функций в режиме реального времени. 6. минимум времени на тестирование.

Прогноз частоты сбоев будем делать на основании предположения, что частота сбоев в будущем будет не меньше, чем частота сбоев в настоящем. Критерием смены прогноза будем считать ситуацию, когда средняя разность между прогнозом и реальностью стала больше некоторого значения. Для принятия решений используются нечеткие числа, обобщающие экспертную информацию, обобщение текущих значений до нечеткого числа и оценка частоты встречаемости значений, оцениваемых одним нечетким числом, на лингвистических шкалах.

7.22. Алгоритм смены прогноза

Рассматривается алгоритм смены прогноза и принятия решений на основе наиболее предпочтительного из «активных процессов» прогноза очередного интервала между сбоями.

1. вычислить формулу методом средних квадратов для некоторого начального участка.

2. проверить совпадение «прогноза» с t_j , моментом очередного сбоя, вычислять уверенность в том, надо ли менять формулу, методом оценки на лингвистической шкале.

3. провести подсчет новых коэффициентов с того момента, когда первый раз окажется, что «прогноз отличается от текущего (реального) значения» (в смысле лингвистических оценок).

4. изменять формулу, если «накопилась уверенность» (в смысле оценки на лингвистической шкале), что ее надо менять.

5. провести новый подсчет коэффициентов, когда «кажется, что старая формула может начать не удовлетворять» (так же в смысле оценки на лингвистической шкале).

6. прогноз вычисляется по нескольким формулам, для каждой формулы вычисляется истинность прогноза $Q = Q + X$, X тем больше, чем больше совпадение значений прогнозируемого момента и действительного момента сбоя. X принадлежит $[-1, +1]$, Q и X вычисляются на лингвистических шкалах. Каждый раз существует «наиболее подходящая» и «наименее подходящая» формулы.

7. исключить формулу, если значения, вычисляемые по ней, начинают «совпадать» со значениями, вычисляемыми по предыдущей формуле. Прогноз или («среднее») совпадает с реальным, если $dt < c_1$, не совпадает, если $dt > c_2$, $[c_1, c_2]$ – отрезок «размытия» понятия «совпадает». Чем ближе к c_1 , тем больше совпадает, чем ближе к c_2 , тем меньше совпадает. «Размытию» соответствуют промежуточные значения. «Совпадает» – никаких действий (если с реальным совпадает только один «процесс» – прогноз). «Не совпадает» подсчет новых коэффициентов (подключение еще одного «процесса» вычисления среднего). На «совпадение» выходит только 1 процесс, остальные отключаются (останавливаются) при приближении к «совпадению». «Приближение» оценивается по лингвистической шкале.

7.23. Оценка правильности прогноза

Оценить «правильность прогноза» можно следующим образом.

Поддерживается несколько процессов прогноза: один из них – «оптимистичный», дает «завышенное время» (в смысле лингвистической оценки по соответствующей шкале), т.е. превышение интервала над реальным; другой – «пессимистичный» – дает «заниженное время», т.е. интервал меньший, чем реальный. Правильность оценивается

следующим образом: «Нормально», если «оптимистичный прогноз» + $c_1 > t_j$ реального, а «пессимистичный прогноз» – $c_2 < t_j$ реального (на «некоторый зазор»). «На пределе» если «пессимистичный прогноз» – c_2 близок («приближается») к реальному t_j («зазор уменьшается»). «Плохо» если «пессимистичный прогноз» – $c_2 > t_j$ реального.

7.24. Отключение процессов прогноза

Чем ближе к текущему моменту время начала «прогноза», тем быстрее он отключается. Более новый «прогноз» останавливается, если его значения «приближаются» к «более старому» прогнозу (т.е. начавшемуся раньше). Среднее (текущее) значение интервала между сбоями, используемое в приведенной выше формуле вычисляется как можно ближе к самому «новому прогнозу» с учетом всех «живущих» прогнозов, причем каждый из них взвешен «оценкой совпадения». На основе этого «текущего среднего» значения изменяется шкала L_t , используемая при оценке работоспособности блока.

7.25. Текущее среднее время

Оценку среднего интервала между сбоями системы без запоминания предыдущей последовательности значений этих интервалов можно осуществлять при помощи накопления информации на лингвистических шкалах и использовать как часть прогнозирующей системы для выбора лучшего прогноза в настоящий момент. Процесс принятия решений о реконфигурации можно строить на основе этой оценки, не используя метод средних квадратов для прогноза очередного интервала между сбоями.

Рассмотрим последовательность интервалов между сбоями в пределе эта последовательность стремится к своему среднему значению:

$$\lim_{i \rightarrow \infty} t_i = t_c$$

Обозначим t_0 среднее значение последовательности, вычисляемое на данный момент времени, t_0 в пределе стремится к t_c . Будем изменять t_0 на величину градации соответствующей шкалы $t_0 = t_0 + V$ или $t_0 = t_0 - V$;

t_0 меняется через определенное число шагов (n), во время которых вычисляется обычное среднее t_{oc}

$$t_{oc}(n) = \frac{\sum_{i=t_0+1}^{t_0+n} t_i}{n}$$

Пусть $dt = t_0 - t_i$, t_0 меняется в соответствии со следующими критериями: dt большое – надо менять t_0 ; dt очень большое – обязательно надо менять t_0 ; dt мало – можно немного изменить t_0 ; dt среднее – надо изменить t_0 ; dt очень мало – можно и не менять t_0 .

Влияние качественных оценок на изменение t_0 учитывается при помощи специальной шкалы L_d , устроенной так же как другие лингвистические шкалы, на которой оценивается dt градациями значений лингвистической пары «большой – маленький». По знаку выражения $dt = t_0 - t_j$ определяется, увеличивать или уменьшать t_0 : «+» – увеличить t_0 , «-» – уменьшить t_0 . Пусть a_t – оценка dt по шкале L_d (большой – маленький), t_0 меняется на величину, связанную с градацией этой шкалы, определяемой значением a_t .

Увеличивать или уменьшать t_0 можно одним из двух способов: а) в зависимости от dt : $t_0 = t_0 + dt * c_i$ – «инерция», т.е. оценка dt по шкале L_d должна быть учтена в c_i ; б) в зависимости от V_t , оценки dt по шкале L_d : $t_0 = t_0 + V_t * c_i$ – где V_t – задает «степень инерции» t_0 относительно t_j , c_i – степень уверенности, что t_0 надо менять.

Более точно, изменение t_0 можно задать следующим образом: $t_0 = t_0 + V_t * a_t * B$, где t_0 – среднее значение бесконечной произвольной последовательности значений, вычисляемое на «данный текущий момент времени» («текущее среднее значение»). V_t – шаг изменения этого значения, определяемый по шкале L_t ; a_t , B – параметры, учитывающие предисторию, определяющие долю шага, на который меняется t_0 : a_t – оценка величины $dt = t_0 - t_j$ на шкале L_d , B – оценка частоты встречаемости a_t – класса.

7.26. Определение работоспособности блока

Принятие решений о реконфигурации можно основывать на t_0 , по методу, описанному выше вычислять значение качественной величи-

ны работоспособности r , применяя t_0 для изменения шкалы L_t , а можно грубо оценить работоспособность блока следующим образом:

Пусть предельное (недопустимое) значение интервала между сбоями для i – того блока, $\min(t_i) = t_{im}$, тогда если $t_i > t_{im}$ – то можно работать, т.е. принимать решение о реконфигурации можно, если «текущее среднее» t_0 приближается (в смысле оценки по соответствующей шкале) к t_{im} . t_{im} определяется при помощи экспертных оценок и, значит, со временем может меняться – уточняться в зависимости от результатов работы системы. Можно поставить задачу прогнозирования возможности возникновения t_{im} .

Все параметры, связанные с качественными величинами в приведенных выше методах, вычисляются на лингвистических шкалах, т.е. получается «много пространственная» картина. Параметры для формулы изменения по выбранной шкале относятся (т.е. изменяются сами) по другим шкалам, т.е. в других пространствах – плоскостях, пересекающихся с «этим параметром» в «точке формулы». Если параметр формулы ни от чего не зависит в данном пространстве ($const$), следовательно, он меняется в других пространствах (некомпланарных к данному) – в формуле учитывается только проекция этого пространства в виде параметра – $const$.

При вычислении среднего значения коэффициенты формул должны меняться по тому же принципу (по шкалам с разными градациями), т.е. выделение классов (большой, маленький и пр.) и оценка количества попаданий в них по жесткому принципу – не подходит, не даст достаточно тонкого результата.

7.27. Изменение значения качественной величины текущего среднего

Изменение значения t_0 в зависимости от t_j можно организовать по тому же принципу как изменение r_i в зависимости от M_j .

Оцениваем по L_d величину $dt = (t_0 - t_j)$. Чем больше dt , тем больше увеличение. Разница между средним и текущим «большая», если «идет превышение среднего над текущим» – тогда уменьшать среднее. Или «постоянно идет превышение текущего над средним» – тогда увеличивать среднее. Чем больше «постоянное превышение», тем больше изменение среднего; в зависимости от «тенденции» изменять

скорость изменения среднего. L_{t+} и L_{t-} – оценка t , следовательно L_d – оценка $t :: =$ взвешивание при помощи лингвистической пары «большой – маленький». Границы L_d определяют «окно». Градации L_d – «различимость» объектов – значений L_t – оценка оценки L_d : взвешивание на весах «часто – редко» L_t – шкала измерения среднего t в зависимости от dt – последней разности $t_c - t_j$. 1. на «единичные» выбросы – не реагировать. 2. чем ближе величина выброса к «постоянному», тем точнее должно быть приближение t_0 к этому значению.

Значение среднего зависит от частоты встречаемости отдельных значений. Чем чаще встречается значение, тем ближе к нему значение среднего. Если увеличивается частота встречаемости другого значения, то среднее сдвигается к этому другому значению. Среднее::= обычное среднее среди «часто встречающихся» значений.

Рассмотренные выше методы дают возможность получить результативную оценку работоспособности сложных систем в тех случаях, когда затруднительно применение точных и вероятностных методов. Предложенные методы и подходы также могут найти применение в других смежных областях [9–13].

Список литературы

1. Хованов Н.В. Математические основы теории шкал измерения качества. Ленинград: ЛГУ, 1982.
2. Нечеткие множества в моделях управления и искусственного интеллекта. Под рук. Пospelova Д.А. М.: Наука, 1986. 396 с.
3. Dubois D., Prade H. Additions of interactive fuzzy numbers. – IEEE Transactions on automatic control, 1981. v. AC-26. № 4.
4. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М., 1976.
5. Балашов Е.П., Куприянов М.С., Логинская Л.Г. Конструирование и интерпретация нечетких алгоритмов // Известия АН СССР. Техническая кибернетика. 1986. № 6. С. 134–139.
6. Куприянов М.С., Логинская Л.Г. Качественный метод представления нечеткой информации / Материалы регионального семинара Надежность технических систем. Ленинград, 1987.
7. Логинская Л.Г. Обеспечение работоспособности вычислительной системы с использованием нечеткого представления качественных понятий. – Второй региональный семинар Распределенная обработка информации, Новосибирск – Улан-Удэ, 1989.

8. Нечеткие множества и теория возможностей. Последние достижения / под ред. Ягера Р.Р. М.: Радио и связь, 1986.

9. *Елизаров С.И.* Адаптивные методы нечёткой кластеризации. Сб. докл.: Международная конференция по Мягким Вычислениям и Измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 2. С. 168–171.

10. *Холод И.И.* Самонастраивающиеся агенты для интеллектуального анализа распределенных данных. Сб. докл.: Международная конференция по Мягким Вычислениям и Измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 1. С. 95–96.

11. *Rukavitsyn A.N., Kupriyanov M.S., Shorov A.V., Petukhov I.V.* Investigation of website classification methods based on data mining techniques // Soft Computing and Measurements (SCM), 2016 XIX IEEE International Conference on. – IEEE, 2016. Pp. 333–336.

12. *Armenak Barsegian, Mikhail Kupriyanov, Ivan Kholod, Sergey Yelizarov, Michael Thess.* Analysis of Data and Processes: From Standard to Realtime Data Mining. Re Di Roma-Verlag, 2014. Pp. 299.

13. *Ivan Kholod, Mikhail Kupriyanov, Andrey Shorov.* «Decomposition of Data Mining Algorithms into Unified Functional Blocks»/ Mathematical Problems in Engineering, 2016. Article ID 8197349, 11 p.

ГЛАВА 8

МЕТОДЫ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ

М.С. Куприянов

8.1. Теория ГА, особенности и основные зависимости

8.1.1. Процесс оптимизации с помощью ГА и понятие схем

Пусть существует первоначальный набор стратегий, сгенерированный случайным образом. Каждая из этих стратегий представляет собой строку длиной L с алфавитом (набором значений в каждой позиции) размером K и является вариантом решения некоторой задачи. Каждая стратегия оценивается по какому-либо критерию; это означает, что каждой стратегии соответствует ее степень пригодности. Можно показать, что оптимальное решение задачи находится в одной из точек пространства поиска K^L [1].

Проиллюстрируем на конкретном примере, каким образом генетический алгоритм направляет поиск в перспективную область пространства поиска на основе информации, доступной на текущем шаге для данной популяции.

Будем считать, что первоначальный набор популяции равен четырём стратегиям (011, 001, 110, 010). Степень пригодности стратегии определяется ее десятичным значением. На первом шаге лучшей является стратегия 110, поскольку она имеет наибольшую степень пригодности. Можно сделать 7 предположений (гипотез), почему эта стратегия имеет лучшие показатели:

- 1) в бите с номером 1 содержится 1 – 1**;
- 2) в бите с номером 2 содержится 1 – *1*;
- 3) в бите с номером 3 содержится 0 – **0;
- 4) в битах с номерами 1 и 2 содержатся 1 и 1 – 11*;
- 5) в битах с номерами 1 и 3 содержатся 1 и 0 – 1*0;

- 6) в битах с номерами 2 и 3 содержатся 1 и 0 – *10;
 7) в битах с номерами 1, 2 и 3 содержатся 1, 1 и 0 – 110.

Схема Н описывает множество точек из пространства поиска задачи, которые имеют некоторые определенные сходства и представляет собой строку, в которой закодирована гипотеза. Порядок схемы $O(H)$ определяется числом однозначно определенных или основных символов алфавита. Если имеется популяция строк длиной L с алфавитом размером K , то схема будет определяться строкой длиной L с алфавитом $K+1$ (* – дополнительный символ алфавита). Таким образом, в рассматриваемом случае существует 27 схем $((K+1)^L, K=2, L=3)$.

Каждая стратегия (строка из пространства поиска) может принадлежать нескольким схемам. Например, стратегии 010 и 110 принадлежат гипотезе *10.

Если $L=3$, то восемь возможных комбинаций бинарных строк и возможные схемы геометрически можно представить в виде куба (рис. 1). Двенадцать ребер куба представляют собой схемы с двумя известными символами в строке и одним неопределенным. На рис. 2 показаны шесть схем, имеющих порядок равный 1, которые представляют собой стороны куба и соответствуют стратегиям с двумя неопределенными символами.

В табл. 1 обобщаются данные, соответствующие порядку, числу и геометрическому представлению схем, а также количеству точек из пространства поиска.

Таблица 1

Представление схем

Порядок схемы $O(H)$	Геометрическое представление	Число стратегий в схеме $2^{L-O(H)}$	Число схем $\binom{L}{O(H)} 2^{O(H)}$
3	Точка	1	8
2	Линия	2	12
1	Плоскость	4	6
0	Куб	8	1
		Всего	27

Предположения о том, почему стратегия 110 является лучшей в популяции были приведены ранее. Аналогичным образом рассмотрим, почему другая стратегия из популяции (010) имеет степень пригодности

равную $2/3$ от средней степени пригодности популяции, в то время как стратегия 110 превышает среднюю степень пригодности в два раза:

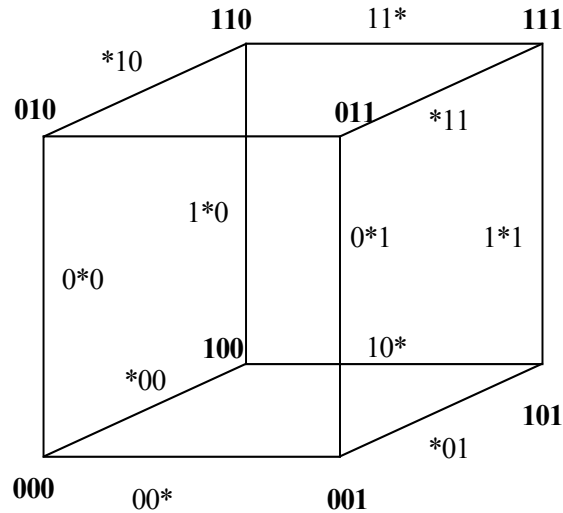


Рис. 1. Геометрическое представление бинарных строк схем

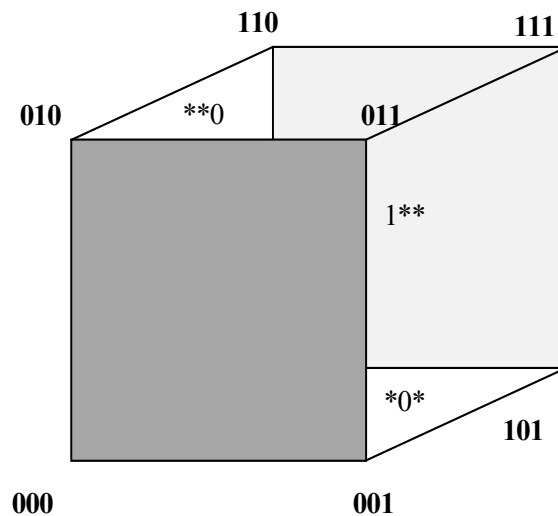


Рис. 2. Геометрическое представление схем с $O(H)=1$

- 1) в бите с номером 1 содержится 0 – 0**;
- 2) в бите с номером 2 содержится 1 – *1*;
- 3) в бите с номером 3 содержится 0 – **0;
- 4) в битах с номерами 1 и 2 содержатся 0 и 1 – 01*;
- 5) в битах с номерами 1 и 3 содержатся 0 и 0 – 0*0;
- 6) в битах с номерами 2 и 3 содержатся 1 и 0 – *10;

7) в битах с номерами 1, 2 и 3 содержатся 0, 1 и 0 – 010.

При этом три из возможных гипотез для двух рассмотренных стратегий конфликтуют друг с другом. Возникновение таких конфликтов может быть объяснено несколькими причинами (нелинейностью задачи, помехами в среде, динамическим поведением среды и т.д.).

В генетическом алгоритме, как и в природе, представление стратегии, ее кодировка является вторичной по степени важности. Первичным является генетический профиль популяции, который содержится в каждой хромосоме, т.е. в каждом бите стратегии.

Все возможные объяснения (гипотезы) и соответствующие им схемы связаны со степенью пригодности каждой из стратегий [2, 3]. Некоторые из этих объяснений без сомнения являются ошибочными, особенно в начальных поколениях ГА.

При переходе от поколения к поколению реализуются два следующих механизма:

- Продолжение поиска в области пространства, в которой можно производить повышение уровня степени пригодности. При этом поиск продолжается в той части пространства, которая состоит из точек с высокой степенью пригодности (или по крайней мере не с низкой степенью пригодности) – исследование пространства.
- Тестирование новых точек из пространства поиска лучших стратегий – эксплуатация пространства.

Таким образом, генетический алгоритм направляет поиск в перспективную область пространства на основе информации, доступной на текущем шаге для данной популяции.

В таблице 2 демонстрируется вхождение схем первоначальных четырех стратегий в поколения 0 и 1. Этим четырем стратегиям, входящим в поколения 0 и 1, соответствуют значения степеней пригодности (табл. 3). Таблицы 4.2 и 4.3 далее используются для иллюстрации последующего процесса формирования нового поколения.

В столбце 3 таблицы 2 содержится количество вхождений $m(H,O)$ для схемы H в поколение 0 популяции. Это число может располагаться в диапазоне от 0 до 4 (что связано с размером популяции M). Сумма этих значений в столбце 3 равняется $32=M2^L$, поскольку каждая стратегия в популяции содержится в восьми гипотезах (включая гипотезу ***) и поэтому требует до восьми вычислений средней степени пригодности схемы. Некоторые гипотезы подкрепляются несколькими стратегиями, а не-

которые не подкрепляются вообще. Столбец 4 показывает среднюю степень пригодности каждой схемы $f(H,O)$.

Таблица 2

Вхождение схем четырех стратегий в поколения 0 и 1

Номер схемы	Схе- ма	Поколение 0		Набор стратегий по- сле селекции (MP)		Поколение 1 после кроссовера	
		m(H,O)	f(H,O)	m(H,MP)	f(H,MP)	m(H,1)	f(H,1)
1	000	0	0	0	0	0	0
2	001	1	1	0-	0	0	0
3	00*	1	1	0-	0	0	0
4	010	1	2	1	2	2+	2
5	011	1	3	1	3	0-	0
6	01*	2	2.5	2	2.5	2	2
7	0*0	1	2	1	2	0-	0
8	0*1	2	2	1-	3	2+	2
9	0**	3	2	2-	2.5	2	2
10	100	0	0	0	0	0	0
11	101	0	0	0	0	0	0
12	10*	0	0	0	0	0	0
13	110	1	6	2+	6	1-	6
14	111	0	0	0	0	1+	7
15	11*	1	6	2+	6	2	6.5
16	1*0	1	6	2+	6	1-	6
17	1*1	0	0	0	0	1+	7
18	1**	1	6	2+	6	2	6.5
19	*00	0	0	0	0	0	0
20	*01	1	1	0-	0	0	0
21	*0*	1	1	0-	0	0	0
22	*10	2	4	3+	4.67	33	3.3
23	*11	1	3	1	3	1	7
24	*1*	3	3.67	4+	4.25	4	4.25
25	**0	2	4	3+	4.67	3	136
26	**1	2	2	1-	3	1	4.2
27	***	4	3	4	4.25	4	16
Всего		32	96	32	136	32	
Среднее			3.00		4.25		
Ненулевых элементов		20	20	16	16	16	

Таблица 3

Вхождение четырех стратегий в поколения 0 и 1

Номер стратегии	Поколение 0			Набор стратегий после селекции		После кроссовера (поколение 1)		
I	Строка X_i	Пригодность $f(X_i)$	Относит. пригодность	Набор стратегий	$f(X_i)$	Точка скрещивания	X_i	$f(X_i)$
1	011	3	0.25	011	3	2	111	7
2	001	1	0.08	110	6	2	010	2
3	110	6	0.50	110	6	-	110	6
4	010	2	0.17	010	2	-	010	2
Всего		12			17			17
Худшая		1			2			2
Средняя		3.00			4.25			4.25
Лучшая		6			6			7

Генетический алгоритм работает только с текущей популяцией определенного размера (в данном случае $M=4$). И только эти четыре стратегии популяции тестируются явным образом. Далее, используя эти четыре стратегии, генетический алгоритм производит селекцию, кроссовер и мутацию с тем, чтобы создать новую популяцию.

Можно видеть, что оценка средней степени пригодности одних гипотез превышает общее среднее значение, а других – ниже общего среднего. Естественно, что поиск должен быть продолжен в пространстве поиска с использованием этого общего среднего значения степени пригодности. Формирование новой популяции должно осуществляться с использованием информации о возможных гипотезах, но при этом доступные гипотезы могут быть ошибочными в том смысле, что всегда существует возможность, что гипотеза, имеющая в данный момент степень пригодности выше среднего, может потерять свою значимость на дальнейшем этапе.

Поэтому вопрос заключается в том, как наилучшим образом использовать текущую доступную информацию. Оптимальным путем формирования новой популяции является размещение имеющихся стратегий с повышением количества будущих вхождений в схему на основе относительной степени пригодности текущей оценки (селекция). Так, например, если текущая оценка средней степени пригодности

сти стратегии превышает в два раза среднюю степень пригодности популяции, то вместо одной позиции в новой популяции она должна занимать две. В контексте рассматриваемого примера необходимо создать новую популяцию из 4 строк длиной 3 бита таким образом, что все 27 возможных схем, к которым принадлежат 4 новые стратегии, одновременно получат новую оценку.

Основополагающая теорема генетических алгоритмов – теорема схем (The Schema Theorem, иначе называемая также теоремой шим) показывает, что генетический алгоритм создает новую популяцию таким образом, чтобы одновременно удовлетворять всем гипотезам.

8.1.2. Теорема схем

Теорема схем подтверждает, что репродукция по пропорциональному принципу (селекция) позволяет формировать следующую популяцию с наилучшими показателями. Необходимые также для процесса эволюции кроссовер и мутация могут как повысить, так и снизить эффективность поиска оптимального решения.

В соответствии с теоремой схем пусть $f(H,t)$ будет средней степенью пригодности схемы H :

$$f(H,t) = \sum_{x_i \in H} f(x_i,t) / m(H,t),$$

где $m(H,t)$ является числом вхождений схемы H в популяцию t .

Теорема схем утверждает, что для генетического алгоритма, использующего операцию репродукции по пропорциональному принципу и генетические операции кроссовера и мутации, ожидаемое число вхождений $m(H,t+1)$ каждой схемы H в популяцию $t+1$ можно определить следующим образом:

$$m(H,t+1) \geq m(H,t) * (f(H,t)/f_{cp}(t)) * (1-\epsilon_c) * (1-\epsilon_m),$$

где ϵ_c – вероятность распада схемы H из-за применения кроссовера; ϵ_m – вероятность распада схемы H из-за применения оператора мутации, $f_{cp}(t)$ – средняя степень пригодности популяции.

Более детально теорема схем иллюстрируется в таблице 2 с использованием данных по четырем стратегиям популяции (табл. 3).

8.1.2.1. Селекция

Таблица 2 показывает эффект оператора селекции на 27 схемах. В этой таблице столбец 5 показывает число вхождений $m(H, MP)$ схемы H в MP (Mating Pool – промежуточный набор стратегий). Знак «+» в этом столбце показывает, что операция репродукции по пропорциональному принципу вызвала увеличение числа вхождений. Такой прирост наблюдается в гипотезах 13, 15, 16, 18, 22, 24 и 25. Следует отметить, что строка 110 принадлежит каждой из этих семи гипотез. Строка имеет степень пригодности, равную 6, что в два раза больше средней степени пригодности, равной 3. В результате выполнения операции селекции данная строка занимает в промежуточном наборе две позиции. Такое копирование соответственно увеличивает количество вхождений в отмеченные схемы и изменяет генетический профиль популяции.

Знак «-» в пятом столбце показывает, что селекция вызвала снижение числа вхождений стратегий в гипотезы в сравнении с предыдущими результатами. Так, стратегия 001 входит во все выделенные таким образом схемы (2, 3, 8, 9, 20, 21, 26) является худшей в поколении 0 и имеет относительную степень пригодности $1/3$. В результате выполнения селекции данная стратегия не копируется в промежуточный набор стратегий.

Выражение $m(H, t+1) = m(H, t) * (f(H, t) / f_{cp}(t))$ описывает эффект селекции. Если предположить, что отдельная схема остается всегда около среднего значения $C * f_{cp}$, где C – константа, то $m(H, t) = m(H, 0) (1 + C)^t$. Данное выражение для поколения $t+1$ популяции относительно поколения t можно преобразовать к виду $m(H, t+1) = m(H, t) * (f_{cp} + C * f_{cp}) / f_{cp} = (1 + C) * m(H, t)$.

Если считать, что популяция, имеющая среднее значение степени пригодности $\bar{f}(t)$ и степень отклонения от этого значения σ , подвергается пропорциональной селекции, то вероятность селекции стратегии n в следующую популяцию может быть выражена как отношение степени пригодности стратегии к суммарной степени пригодности популяции: $P_{select}(n) = f(n) / \sum f(k)$.

Среднее значение степени пригодности следующей популяции может быть, в свою очередь, представлено в [4] как $\bar{f}'(t) = \bar{f}(t) + S(t)$, где $S(t)$ представляет собой дифференциал селекции и определяется как разность между средней степенью пригодности выбранных родителей $f_{cps}(t)$ и средним значением степени пригодности популяции: $S(t) = f_{cps}(t) - f_{cp}(t)$.

$S(t)$ может быть выражено также через степень отклонения σ (СКО) как: $S(t) = \frac{\sigma^2(t)}{\bar{f}(t)}$.

Таким образом, улучшение в результате селекции средней степени пригодности зависит от отклонения степени пригодности в популяции. В то же время, селекция имеет влияние на сокращение степени отклонения в популяции (сходимость), следовательно, она будет производить только ограниченное повышение эффективности, поскольку отклонение может снизиться до нуля.

8.1.2.2. Кроссовер

Столбец 7 таблицы 2 отражает количество вхождений в каждую из 27 гипотез, а столбец 8 – среднюю степень пригодности каждой гипотезы после применения селекции и кроссовера в поколении 1.

Знак «+» в столбце 7 показывает, что кроссовер вызвал увеличение числа вхождений в гипотезы по сравнению с предыдущим шагом (столбец 5). Это можно увидеть в строках таблицы с номерами 4, 8, 14, 17. При этом видно, что строки 14 и 17 не были представлены в популяции ранее. А схема 14 представляет собой оптимальное решение задачи.

Генетическая операция кроссовера является необходимой функцией создания новых стратегий в пространстве поиска, однако, как отмечалось ранее, она способна как усилить, так и ослабить популяцию.

Для обычных операций генетических алгоритмов со строками определенная длина $\delta(H)$ схемы H является расстоянием между крайними значащими символами, т.е. не *. Число точек скрещивания при этом колеблется в диапазоне $L-1$. Например, $\delta(1*1)=2$, а $\delta(*11)=1$. Если точка скрещивания будет находиться между первой и второй позициями строки, то схема *11 не будет разрушена кроссовером. Если точка скрещивания находится внутри схемы, и вторая родительская строка, принимающая участие в скрещивании, не принадлежит этой схеме, то строка наследника скорее всего не будет принадлежать этой схеме.

Вероятность такого разрушения схемы при кроссовере: $\varepsilon_c = \delta(H)/L-1$. Поэтому ε_c будет небольшой, если $\delta(H)$ небольшая.

8.1.2.3. Мутация

Генетический оператор мутации позволяет вносить изменения в популяцию. Такое изменение способно оказывать как положительное, так и отрицательное влияние на популяцию, приводя к формированию

индивидов как с высокой, так и низкой степенью пригодности. Отрицательное влияние на популяцию будет небольшим для схем с низкой степенью определенности (т.е. с большим количеством символов *). Например, случайная мутация строки 011 имеет больший шанс принадлежать схеме **1, чем схеме *11.

Таким образом, вероятность такого разрушения схемы при операции мутации может быть представлена выражением $\epsilon_m = (1 - p_m)^{O(H)} \approx 1 - p_m O(H)$, где $O(H)$ порядок схемы. Поэтому ϵ_m будет небольшой, если $O(H)$ небольшой.

Таким образом, размещение стратегий находится недалеко от оптимума, когда вероятности разрушения схемы при кроссовере и мутации небольшие. Схема с относительно короткой определенной длиной, высокой приспособленностью (степень пригодности) и низким порядком (строющий блок [5]) будет переходить из популяции к популяции, приближаясь к оптимуму.

В заключение следует отметить, что генетический алгоритм имеет свойство неявного параллелизма, которое позволяет создавать строки новой популяции таким образом, что схемы, представляющие эти строки, могут быть все автоматически представлены пропорционально отношению степени пригодности схемы $f(H, t)$ к средней степени пригодности популяции $f_{cp}(t)$.

8.1.3. Повышение эффективности генетических алгоритмов

Теорема схем является базовым теоретическим обоснованием ГА и раскрывает основной принцип работы ГА (пропорциональную селекцию), который обуславливает продвижение к нахождению оптимального решения. Однако она не позволяет предсказать, когда ГА приблизится к оптимуму и насколько хорошее решение будет найдено.

Основным, согласно теореме схем, является механизм селекции, мутация и кроссовер представляются как побочный механизм и практически не учитываются. Теорема Схем не выявляет те компоненты поиска ГА, от которых зависит эффективность, и не дает возможность отличить эффективные ГА от неэффективных. Она утверждает, что схемы со степенью пригодности выше средней (особенно короткие, низкого порядка схемы – строящиеся блоки), экспоненциально увеличивают свою частоту в популяции каждое поколение. Ошибка заключается в том, что такой рост схем не влияет на качество поиска, выполненного ГА.

Значение теоремы схем заключается в том, что ГА сосредотачивает поиск на перспективных областях пространства, и таким образом увеличивает вероятность того, что новые выборки из пространства поиска будут иметь более высокую пригодность. Однако можно только предположить, что области – это схемы, а перспективными областями являются схемы со степенью пригодности выше средней.

Наследник, произведенный в результате кроссовера, скорее всего будет получен из той же самой «области», что и родители, но это зависит от уровня разрушения при кроссовере, в теореме схем такая корреляция не учитывается.

В 1975 году Холландом (Holland) была предложена каноническая модель ГА для произвольных схемы представления (кодировки), операторов и функций пригодности:

$$p(x)' = \sum_{y,z \in S} T(x \leftarrow y, z) \frac{\omega(y)\omega(z)}{\bar{\omega}^2} p(y)p(z),$$

где: $p(x)$ – частота индивида x (количество индивидов x) в популяции;

$p'(x)$ – частота индивида x в следующем поколении;

$\omega(x)$ – степень пригодности индивида x ;

$\bar{\omega}$ – среднее значение степени пригодности популяции $\sum_x \omega(x)p(x)$;

S – пространство поиска из n индивидов;

$T(x \leftarrow y, z)$ – передаточная функция, вероятность того, что наследник с генотипом x производится родителями с генотипами y и z как результат действия генетических операторов на заданной схеме представления, при этом $T(x \leftarrow y, z) = T(x \leftarrow z, y)$ и $\sum_x T(x \leftarrow y, z) = 1$ для всех $y, z \in S$.

Передаточная функция должна нести в себе информацию о структуре хромосом индивидов и генетических операторах, что связано с динамикой ГА.

Согласно теореме Прайса [6, 7], эффективность поиска решения зависит от величины ковариации между функцией измерения и степенью пригодности. В общем случае функция измерения $F(x)$ представляет собой некоторое свойство генотипа x с $F: S \rightarrow V$, где V – пространство векторов вещественных чисел. Среднее значение функции измерения в текущем и следующем поколениях может быть представлено как:

$$\bar{F} = \sum_x F(x)p(x), \quad \bar{F}' = \sum_x F(x)p(x)',$$

для каждой родительской пары $\{y, z\}$ $\phi(y, z)$ представляет собой ожидаемое значение функции измерения наследника:

$$\phi(y, z) = \sum_x F(x)T(x \leftarrow y, z)$$

Среднее значение функции измерения в следующем поколении:

$$\bar{F}' = \bar{\phi} + Cov[\phi(y, z), \omega(y)\omega(z) / \bar{\omega}^2]$$

где:

$\bar{\phi} = \sum_{y, z} \phi(y, z)p(y)p(z)$ – среднее ожидаемое значение функции измерения наследника,

$$Cov[\phi(y, z), \omega(y)\omega(z) / \bar{\omega}^2] = \sum_{y, z} \phi(y, z) \frac{\omega(y)\omega(z)}{\bar{\omega}^2} p(y)p(z) - \bar{\phi}$$

Выбор одной функции измерения (функции частоты схемы H – $F(x, H)=1$, если $x \in H$ и 0 в противном случае, которая выделяет частоту схемы из популяции) дает Теорему Схем, в то время как выбор другой функции измерения (функции измерения – $F(x, \omega)=1$, если $\omega(x) > \bar{\omega}$ и 0 в противном случае) выделяет распределение степени пригодности из популяции для оценки эффективности.

Распределение степени пригодности считается основой оценки эффективности ГА при сравнении способности ГА генерировать нового, высоко приспособленного индивида с эффективностью произвольного поиска.

Основным результатом исследований [8] явилось то, что для ГА, лучшего, чем произвольный поиск, должна иметься корреляция между степенью пригодности родителей и распределением степени пригодности их наследника.

Полученные в [9] выражения для корреляции родителей и наследников при использовании операторов кроссовера включают в

себя понятие схем. При этом для эффективного ГА должна иметься корреляции между степенью пригодности схемы и степенью пригодности наследника.

Этот факт получил свое практическое подтверждение. Для доказательства этого факта было рассмотрено двоичное и десятичное представление решения задачи в строке индивида. За основу был взят БГА, использующий только селекцию и однородный кроссовер. Тестирование БГА проходило с усреднением по 500 запускам и числе поколений, равном 100 для каждого запуска. Для оценки алгоритмов использовались функции on-line и off-line, определенные в [10]:

$$\begin{aligned} onLine(T) &= 1/Tn \sum_{t=1}^T \sum_{i=1}^n f(t,i), \\ offLine(T) &= 1/T \sum_{t=1}^T f(t^*), \end{aligned}$$

где $f(t,i)$ – значение функции степени пригодности индивида i в популяции t , $f(t^*)$ – лучшее значение в поколении t .

Результаты теста, представленные на рис. 2.3, доказывают влияние способа представления решения задачи на эффективность кроссовера и БГА в целом.

Способ представления решения задачи с помощью двоичной и десятичной кодировки в свою очередь связан с числом возможных схем. Рассмотрим, например, 20-битовую строку. Эта строка может представлять 2^{20} возможных решений задачи и дает 3^{20} различных схем. Число 2^{20} соответствует примерно числу 10^6 . Следовательно, решение задачи можно представить также в виде строки из шести десятичных чисел. Соответственно число различных схем – 11^6 , что приблизительно в 2000 раз меньше, чем в случае двоичной кодировки. Таким образом, увеличение числа возможных схем приводит к повышению эффективности кроссовера и БГА в целом.

Кроссовер (рекомбинация) может быть представлен и выполнен в виде маски вектора $r \in \{0,1\}^L$ двоичных переменных $r_i \in \{0,1\}$, которые обозначают локусы (см п. 8.1.2.1), передаваемые от любого родителя. Так, все локусы с $r_i=0$ передаются от одного родителя, в то время как локусы с $r_i=1$ передаются от другого родителя. Векторы $r=0=(0\dots 0)$ и $r=1=(1\dots 1)$ соответствуют отсутствию рекомбинации. Действие любого оператора кроссовера может быть представлено как распределение вероятности $R(r)$ на наборе $r \in \{0,1\}^L$.

В [9] отмечается, что однородная рекомбинация является лучшей стратегией изменения $R(r)$.

Таким образом, имеется теоретическое обоснование и практическое подтверждение существенного влияния кроссовера, а также его типа на эффективность ГА.

8.1.4. Разработка обобщенной схемы ГА

ГА является динамическим процессом, который состоит из трех частей: оценки степени пригодности популяции, селекции и преобразования (или разрушения), который является неминуемым результатом использования генетических операторов. Применение генетических операторов приводит к появлению нового генетического материала. Это определяет дальнейшее направление исследований, целью которых является выявление способов повышения эффективности ГА.

ГА могут иметь большое число модификаций и могут отличаться:

- 1) генетическим представлением задачи;
- 2) механизмом оценки индивидов (функцией степени пригодности);
- 3) способом селекции и связанными числом отобранных родителей для выполнения одной итерации ГА, числом наследников, формируемых в результате одной итерации ГА;
- 4) схемой применения основных генетических операторов (кроссовера, мутации и простой репродукции, т.е. селекции без последующего применения кроссовера и мутации);
- 5) типом генетических операторов.

Составляющие 1 и 2 вместе с таким параметром как размер популяции обуславливаются решаемой задачей и соответственно не могут быть определены однозначно для произвольного ГА. Наибольший интерес в этом смысле представляют составляющие 3-5. В приложении 2 приводятся результаты исследований с целью определения эффективности этих составляющих. При этом эффективность ГА существенно зависит от таких взаимосвязанных параметров как вероятность применения и тип генетических операторов, размер популяции.

Для исследования влияния составляющих БГА преобразован в обобщенный БГА, имеющий более гибкую структуру. Обобщенный БГА представлен на рис. 4 и 5.

Особенностью обобщенного алгоритма является наличие двух схем, определяющих последовательность применения генетических операторов:

1. Вероятность репродукции (P_R) и вероятность кроссовера (P_C) в сумме равны 100%. Если кроссовер выполняется, то с вероятностью мутации (P_M) к результату, полученному от кроссовера, применяется мутация.

2. Вероятность репродукции (P_R), вероятность кроссовера (P_C) и вероятность мутации (P'_M) в сумме равны 100%. В этом случае операции применяются к разным индивидам.

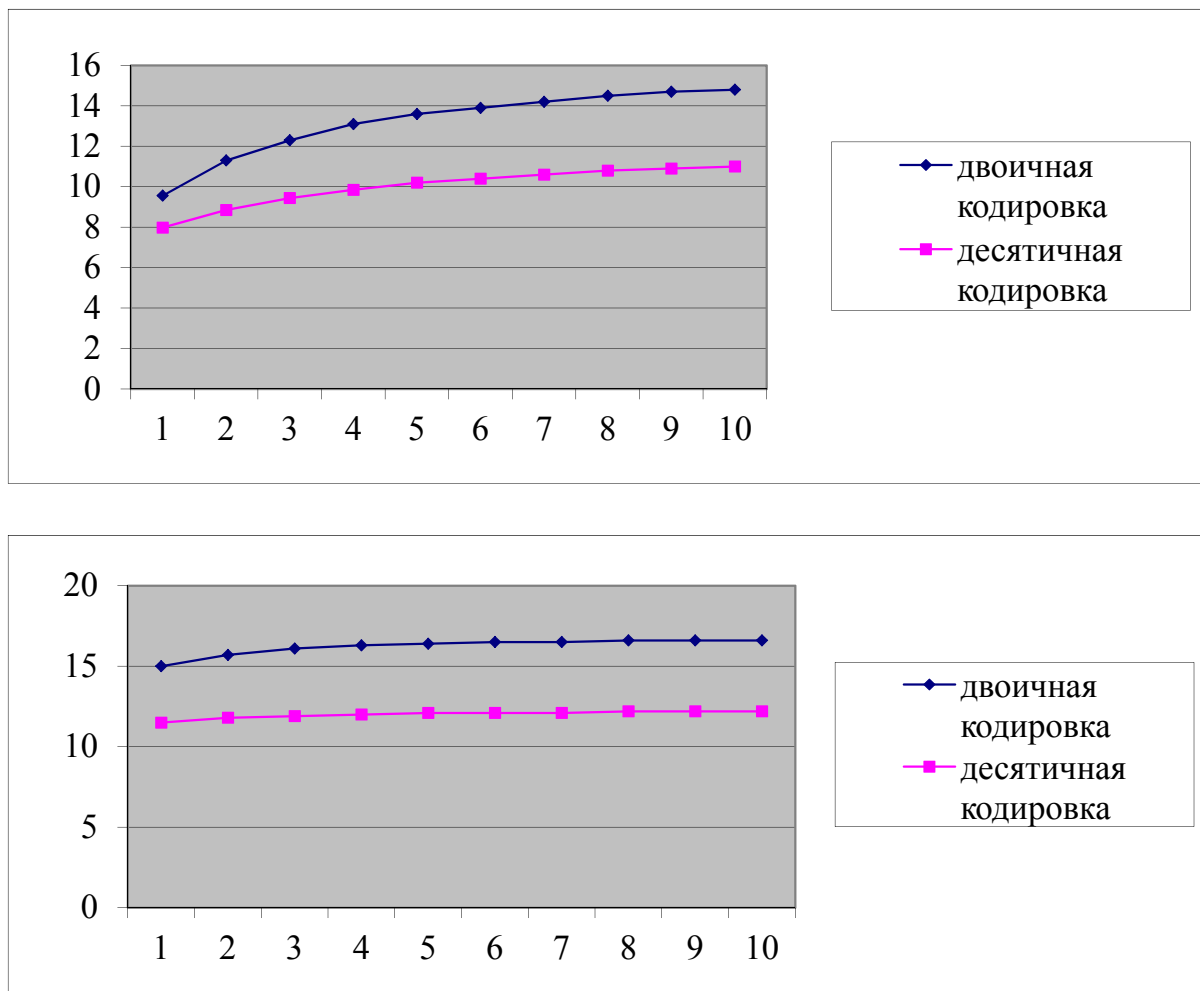


Рис. 3. Влияние способа представления решения задачи на эффективность кроссовера. P – число поколений БГА

Вероятность кроссовера и репродукции определяют процент индивидов, к которым применяются эти операторы и однозначно определяются параметром P_C и P_R . Вероятность мутации P'_M задает процент индивидов, к которым применяется мутация. Вероятность мутации P_M , используемая в ГА как параметр мутации, при этом определяется как

$P'_m \cdot 1/n$, где n – длина строки индивида, $1/n$ – интенсивность мутации. Таким образом, на вероятность мутации влияет длина строки.

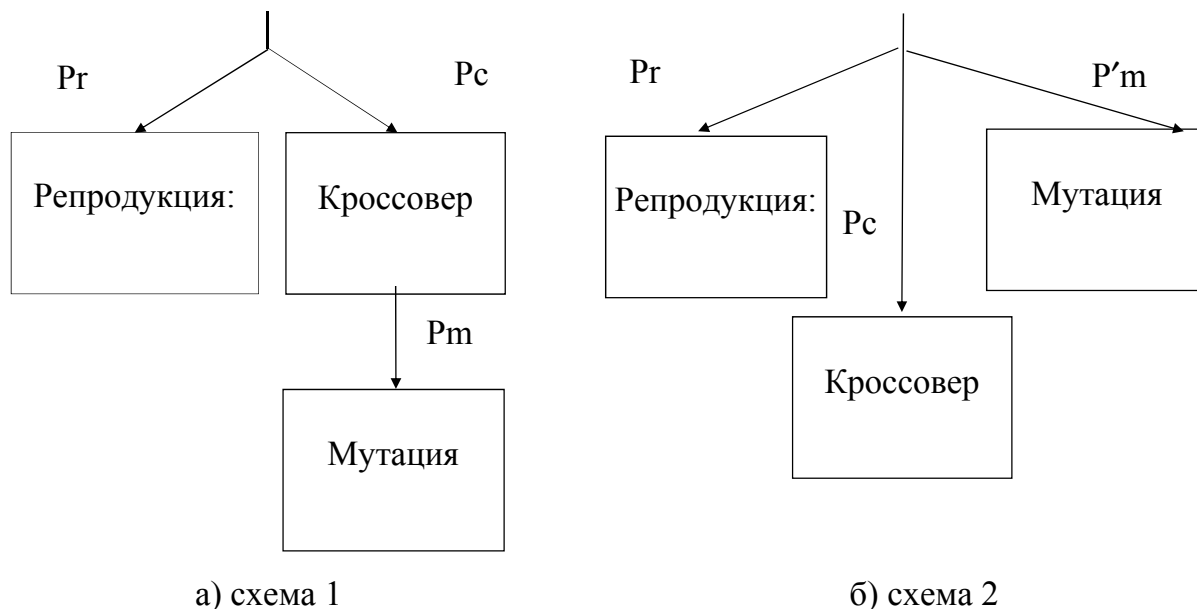


Рис. 4. Схемы 1 и 2 обобщенного БГА

Параметр T определяет процент лучших индивидов популяции, которые будут отобраны для применения генетических операторов. В этом случае $T\%$ индивидов будут однозначно переведены (репродуцированы) в следующую популяцию без изменений.

Другой особенностью обобщенного ГА является вероятностный выбор одного или двух наследников, произведенных родительской парой. В принципе, возможно применение различных комбинаций числа родителей и наследников для выполнения кроссовера в диапазоне от 2 до N , где N – размер популяции. Может быть предложено большое число механизмов кроссовера, включая N -точечный или однородный. Однако увеличение числа родителей увеличивает вероятность разрушения схем в случае N -точечного кроссовера, т.к. наследники формируются из частей, число которых равно числу родителей. Однородный кроссовер нескольких родителей в одной точке будет сведен к вероятностному выбору двух родителей. Возможен также вариант, когда число родителей меньше числа наследников. Этот вариант соответствует случаю, когда при постоянном размере популяции часть популяции резервируется для дополнительных наследников, тем самым, уменьшая число отобранных родителей.

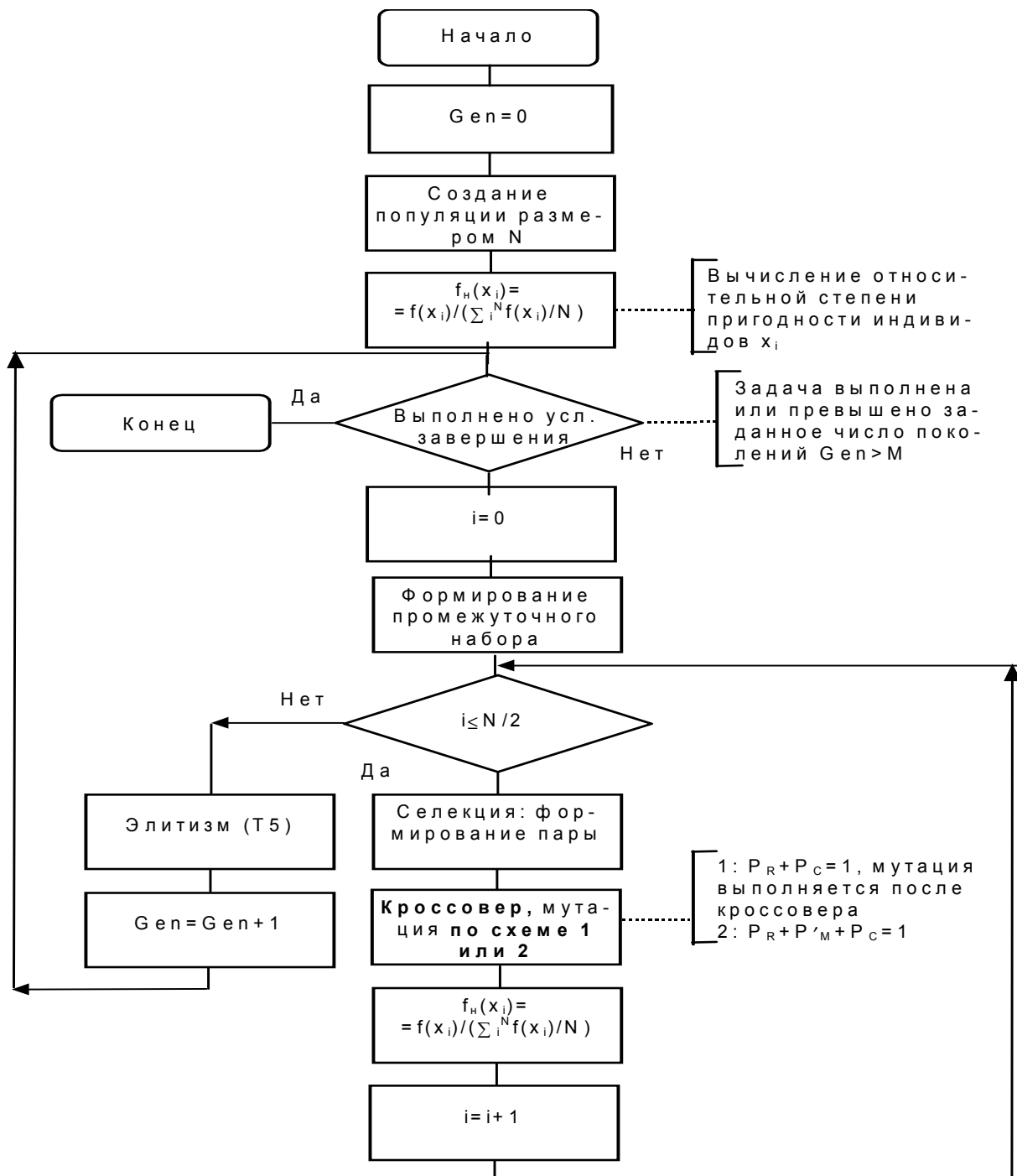


Рис. 5. Обобщенный БГА

8.2. Исследование и разработка оператора селекции

Существуют модификации ГА, отличающиеся от стандартного прежде всего механизмом селекции – ГАС (массовая селекция или се-

лекция округления) и SSGA. В случае массовой селекции задается пороговое значение $Trunc\%$ – процентное соотношение лучших индивидов, которые будут выбраны в качестве родителей. Обычно $Trunc$ устанавливается в диапазоне от 10% до 50%. SSGA предполагают наиболее простой механизм селекции. В этом случае одновременно необходимо найти двух лучших и худшего индивидов. К лучшим применить кроссовер и мутацию. Худшего индивида заменить наследником. При этом возможна также сравнительная оценка степени пригодности наследника и худшего индивида. Тогда, в случае, если степень пригодности наследника хуже, замещение не производится.

Элитизм является оператором селекции, который комбинируется с другими операторами. При этом наилучший индивид или несколько индивидов автоматически копируются в следующую генерацию без случайного отбора. Это гарантирует, что лучший индивид не будет потерян в следующей популяции. Многими авторами во многих публикациях признается важность использования данного типа селекции. Интуитивно понятно, что использование данного оператора положительно скажется на эффективности ГА. Его применение имеет смысл, если временной критерий работы ГА не учитывается. Вместе с тем этот механизм прост с точки зрения реализации и несущественно повысит вычислительную сложность ГА.

В разделе предлагается и исследуется новый оператор селекции, используемый в БГА так, как это представлено на рис. 6.

БГА с предлагаемым оператором селекции (БГА+) находится на стыке БГА, оперирующего популяцией и формирующего промежуточный набор описанными выше способами, и SSGA, который оперирует парой индивидов. Особенность предлагаемого оператора селекции заключается в том, что он позволяет избежать формирования промежуточного набора в отличие от БГА, это приводит к экономии памяти для хранения промежуточного набора, а также позволяет сохранить принцип пропорциональности в отличие от SSGA.

Отличие БГА+ от БГА проиллюстрировано на рис. 7. От SSGA БГА+ отличает разный подход к времени жизни индивидов. В SSGA одновременно в популяции существуют родительская пара и наследник. При этом родительская пара с наилучшей степенью пригодности может находиться в популяции до окончания работы алгоритма. Следовательно, в SSGA время жизни родительской пары не используется.

БГА с модифицированным оператором селекции (БГА+).

Шаг 0: Определение генетического представления задачи.

Шаг 1: Создание первоначальной популяции $P(0)=x_1^0, \dots, x_N^0, t=0$.

Шаг 2: Вычисление *относительной (нормализованной) степени пригодности* $f_n(x_i) = f(x_i) / \sum_i^N f(x_i) / N$.

Шаг 3: Выбор пары из лучших индивидов. Выбор пары из худших индивидов.

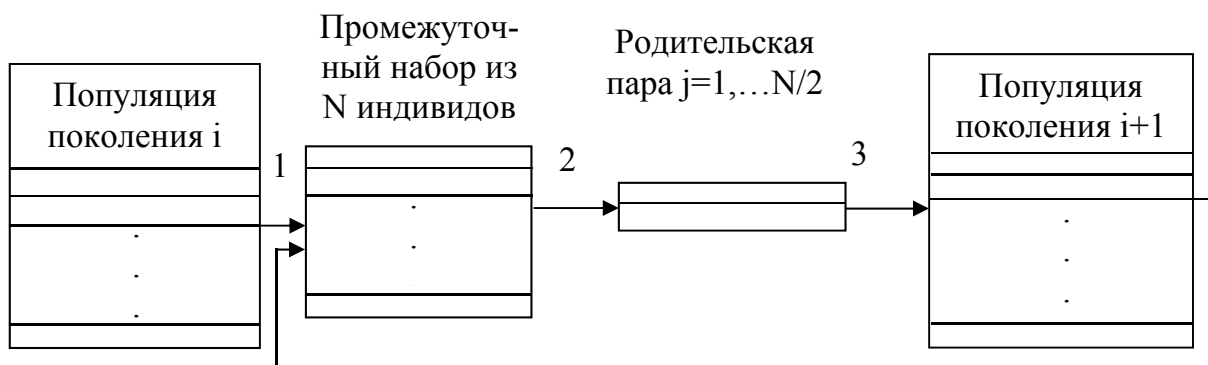
Применение мутации и (или) кроссовера к выбранной паре лучших индивидов.

Результаты замещают пару худших индивидов, которая далее в процессе селекции данной популяции не участвует, формируя новую популяцию $P(t+1)$. Вероятность участия лучших индивидов в дальнейшем процессе селекции популяции снижается на нормализованное среднее значение степени пригодности $f_{cp} / \sum_i^N f(x_i)$.

Шаг 4: Число сформированных пар увеличивается на 1. Возврат к шагу 3 продолжается до $N/2$.

Шаг 5: $t=t+1$, возврат к шагу 2.

Рис. 6. БГА+



а) БГА



б) БГА+

- 1 – формирование промежуточного набора
- 2 – формирование родительской пары
- 3 – применение генетических операторов

Рис. 7. Формирование поколения i+1 в БГА+ и БГА

В БГА+ все индивиды находятся в равных условиях, время жизни родительской пары ограничено одним поколением. БГА+ предполагает на каждом шаге отбор наилучших индивидов популяции в качестве родительской пары аналогично SSGA. Такой отбор представляется более обоснованным с точки зрения теоремы схем, чем произвольный отбор родительской пары БГА

8.3. Исследование операторов мутации и кроссовера

Помимо основных типов кроссовера (одноточечного, двухточечного, однородного, кроссовера с однобитовой адаптацией), существуют также:

- Кроссовер с переносом позиций битов двух строк [11].
- Рекомбинация генетического набора (Gene Pool Recombination, Н. Muhlenbein, Н.-М. Voight), из строк заданного числа родителей, число может быть больше двух, случайным образом выбираются гены.
- Кроссовер с несколькими родителями для формирования одного наследника [12].
- Hill-climbing кроссовер, представляющий собой комбинацию одного из перечисленных стандартных типов кроссовера и локального механизма Hill-climbing [13].
- Нечеткий кроссовер [14], для которого предложены различные типы кроссовера, базирующиеся на нечетких соединителях: компенсационных операторах, t-norms, t-conorms и т.д. Применительно к этим операторам кроссовера был предложен целый ряд механизмов селекции наследников (OSM), который позволяет избежать преждевременной сходимости популяции.

Результаты исследований основных типов мутации, кроссовера, а также размера популяции приводятся в **приложениях 2.2–2.4**. Результаты тестов БГА+ позволяют сделать следующие выводы:

1. При использовании в БГА+ мутации с вероятностью 0,01 БГА+ обеспечивает лучшее по качеству решение задачи, оцениваемое по функциям onLine и offLine. В данном тесте значение вероятности кроссовера устанавливалось в 0, размер популяции задавался значением 60, а число поколений – 100. При тестировании использовались следующие значения вероятности мутации: 0,01, 0,005 и 0,002.

2. При использовании в БГА+ оператора кроссовера с вероятностью 0,9 БГА+ обеспечивает лучшее по качеству решение задачи, оце-

ниваемое по функциям onLine и offLine. В данном тесте значение вероятности мутации устанавливалось в 0, размер популяции задавался значением 60, а число поколений – 100. При тестировании использовались следующие значения вероятности кроссовера: 0,9, 0,5 и 0,2.

3. При задании в БГА+ значения размера популяции, равного 60, БГА+ обеспечивает лучшее по качеству решение задачи, оцениваемое по функциям onLine и offLine. Тестирование проходило при следующих значениях параметров ГА: вероятность мутации – 0,01, вероятность кроссовера – 0,9, число поколений – 100, при этом порядок применения операторов кроссовера и мутации определялся схемой 1 (применение мутации к индивидам, уже подвергшимся кроссоверу). При тестировании использовались следующие значения размера популяции: 60, 40 и 20.

4. При использовании в БГА+ оператора однородного кроссовера, как в случае использования оператора мутации, так и без него БГА+ обеспечивает лучшее по качеству решение задачи, оцениваемое по функциям onLine и offLine. Тестирование проходило при следующих значениях параметров ГА: вероятность мутации – 0,01 или 0, вероятность кроссовера – 0,9, число поколений – 100. В случае использования и оператора кроссовера, и оператора мутации порядок применения операторов кроссовера и мутации также определялся схемой 1. При тестировании использовались следующие типы оператора кроссовера: одноточечный, двухточечный, однородный, однобитовый локальный и однобитовый глобальный.

Для оценки алгоритмов использовались функции onLine и offLine, определенные в [21]:

$$onLine(T) = 1/Tn \sum_{t=1}^T \sum_{i=1}^n f(t,i),$$

$$offLine(T) = 1/T \sum_{t=1}^T f(t^*),$$

где $f(t,i)$ – значение функции степени пригодности индивида i в популяции t , $f(t^*)$ – лучшее значение в поколении t .

8.4. Разработка нечеткого адаптивного однородного кроссовера

8.4.1. Сущность и способы реализации

Теоретические исследования доказывают существование «эффекта кроссовера» как вероятности разрушения схем и генетического

представления решения задачи (строки индивида), а также влияния кроссовера и его типа на эффективность ГА (подраздел 8.1).

Результаты исследований, приведенных в приложении 2, позволяют сделать следующие выводы:

- Генетический оператор однородного кроссовера является наиболее эффективным среди других стандартных типов (неадаптивных) операторов кроссовера.

- С кроссовером связана проблема преждевременной сходимости ГА, которая требует решения, при этом должно быть установлено разумное сочетание интенсивности или вероятности применения кроссовера и мутации.

- Вероятность применения генетического оператора влияет на эффективность ГА, при этом степень и характер этого влияния определяется ландшафтом степени пригодности популяции и уровнем развития популяции. В диссертационной работе предлагается рассматривать отдельно понятия вероятности и интенсивности генетических операторов мутации и однородного кроссовера. Интенсивность генетических операторов мутации и однородного кроссовера определяется как d/n , где d – число битов строки, к которому применяется генетический оператор, а n – длина строки индивида. Для оператора мутации $d=1$.

- Адаптивный метод (однобитовая адаптация) позволяет делать структуру ГА более гибкой, тем самым повышая эффективность алгоритма.

Помимо однобитовой адаптации может быть реализовано большое количество вариантов комбинированных (использующих мутацию и кроссовер) адаптивных (меняющих вероятность или интенсивность операторов) ГА.

В процессе работы алгоритма на основании данных о ландшафте степени пригодности текущей популяции должны происходить изменения вероятности или интенсивности применения генетических операторов, которая должна уменьшаться по мере приближения к оптимуму с тем, чтобы снизить эффект разрушения. В данном алгоритме закладывается модель принятия решения, предполагающая задание зависимостей параметров ГА, влияющих на качество решения задачи. Информация по зависимостям параметров ГА обладает элементами нечеткости и неопределенности, определяемой адаптивным (зависящим от текущей ситуации) характером работы ГА. Решение принимается на основе этой нечеткой информации. Таким образом, нечеткая

база правил (НБП) способна определять работу ГА, внося элемент адаптации под меняющиеся условия. Каждое правило содержит в antecedентной части переменные, связанные с характером ландшафта степени пригодности в популяции, а в части решения содержит переменные, управляющие вероятностью или интенсивностью применения генетических операторов.

В разделе предлагаются и анализируются два варианта реализации нечеткого адаптивного однородного кроссовера (НАОК 1 и НАОК 2), в которых НБП управляет интенсивностью однородного кроссовера. Целью использования НАОК является повышение эффективности ГА. Эффективность ГА определяется по оценочным функциям onLine и offLine, позволяющим получать зависимости значения степени пригодности популяции от числа поколений (или шагов выполнения ГА).

НАОК 1 и НАОК 2 отличаются принципом формирования НБП. НАОК 1 использует одну директивно заданную на первом шаге работы ГА НБП. В отличие от НАОК 1 НАОК 2 предполагает наличие $N/2$ НБП, где N – размер популяции ГА, а также настройку НБП в режиме on-line за счет введения ГА верхнего уровня. При этом популяция НБП эволюционирует вместе с популяцией строк индивидов, которые представляют собой решения задачи, с помощью собственного процесса селекции, кроссовера и мутации. Данный способ исключает необходимость директивного задания исходной НБП.

Критерий качества или степень пригодности нечетких правил вычисляется из степени пригодности стратегий нижнего уровня, представляющих собой решения тестовой задачи с помощью ГА.

Однозначно определяется, что при использовании ГА с НАОК вычислительная сложность растет, при этом вычислительная сложность НАОК 2 выше вычислительной сложности НАОК 1. В случае использования последовательного БГА это приводит к увеличению времени работы ГА. Эта проблема может быть решена за счет перехода от последовательного ГА с НАОК к параллельному ГА с НАОК.

8.4.2. Разработка НБП для НАОК

Два типа НАОК (НАОК 1 и НАОК 2), используемых в модифицированном базовом ГА (БГА+), проиллюстрированы на рис. 4.8. НАОК применяется к родительской паре популяции размером N .

В общем случае НБП может быть сформирована для управления и адаптации другими параметрами ГА (см. п. 8.2.5). В ГА, основанном на

НАОК, НБП содержит правила, в которых областью определения значений двух входных переменных А и В являются значения показателей эффективности родителей (f_A , f_B), а выходная переменная С определена на значениях интенсивности I_C однородного кроссовера. При этом число значений входных переменных одинаково. В результате нечеткого логического вывода на выходе формируется значение интенсивности однородного кроссовера (I_C) в соответствии со значением выходной переменной С. Для каждой пары индивидов значение интенсивности однородного кроссовера определяется процентным отношением d/n , где d – число битов строки, к которому применяется генетический оператор, а n – длина строки индивида. Число битов, подвергающихся кроссоверу, может лежать в диапазоне от 0 до $n/2$, где n – длина строки индивида, интенсивность однородного кроссовера может лежать в диапазоне 0-50%. Ниже рассмотрены компоненты и способы настройки НБП.

1. Входы и выходы системы

НАОК предполагает адаптацию на уровне индивидов, а не популяции. Это приводит к тому, что входными переменными являются не показатели эффективности популяции (как это было рассмотрено в п.1.2.5), а показатели эффективности родителей (родительской пары). Для БГА+НАОК I_C является выходной переменной, входными переменными являются следующие разработанные показатели эффективности родительской пары:

- $PDM1' = f_i / f_{\max}$, где $PDM1' \in [0, 1]$, f_i – значение степени пригодности индивида $i \in [1, N]$, $f_{\max}$ – максимальное значение степени пригодности текущей популяции;
- $PDM1'' = f_i / f_{\max}$, где $PDM1'' \in [0, 1]$, f_i – значение степени пригодности индивида $i \in [1, N]$, $f_{\max}$ – требуемое значение степени пригодности (критерий качества).

$PDM1'$ является динамическим показателем, т.к. шкала является динамической, поскольку это эквивалентно $PDM1' = f_i$ при диапазоне значений $[0, f_{\max}]$. $PDM1''$ является статическим, т.к. в этом случае может быть использовано непосредственное значение степени пригодности на фиксированной шкале $[0, f_{\max}]$.

2. Лингвистические термы и функции принадлежности

Каждый вход и выход должен иметь набор лингвистических термов, определяющихся через функции принадлежности нечетких множеств. Функции принадлежности лингвистических термов для двух входных и одной выходной переменной НАОК заданы так, как пока-

зано на рис. 9. Смежные функции принадлежности накладываются друг на друга так, что центр предыдущей функции принадлежности совпадает с крайней отметкой следующих. Такой способ задания обеспечивает экономию памяти при реализации.

3. Нечеткие правила

После выбора входов и выходов и определения лингвистических термов и функций принадлежности определим принцип формирования нечеткой базы правил, описывающий отношения между термами (см. табл. 4).

Таблица 4²

База правил

Входы		Выход
A	B	C
Low	Low	High
Low	Medium	High
Low	High	Medium
Medium	Low	Medium
Medium	Medium	Medium
Medium	High	Low
High	Low	Medium
High	Medium	Low
High	High	Low

НБП, используемые в НАОК, имеют фиксированный набор правил определяющийся полным перебором сочетаний лингвистических термов входных переменных, число правил определяется выражением $N_A \times N_B$. Однотипность НБП необходима для использования ГА верхнего уровня с целью адаптации НБП. Одинаковая длина строк, каждая из которых представляет собой одну НБП, упрощает выполнение кроссовера, исключая необходимость выравнивания строк. В рассматриваемом случае используются НБП из девяти правил.

Обобщая вышесказанное, взаимодействие ГА и НПА в НАОК 1 и НАОК 2 выражается в наличии следующих блоков (см. рис. 8), где: F0 –

² Таблица 4. Содержит базу правил, отражающую экспериментально исследованные в Главе зависимости интенсивности кроссовера от показателей эффективности родительских индивидов. База правил используется в НАООС1 для получения экспериментальных результатов в п. 8.4.4

блок генерации; F1 – блок нечеткого логического вывода; F2 – адаптация популяции НБП с помощью ГА верхнего уровня.

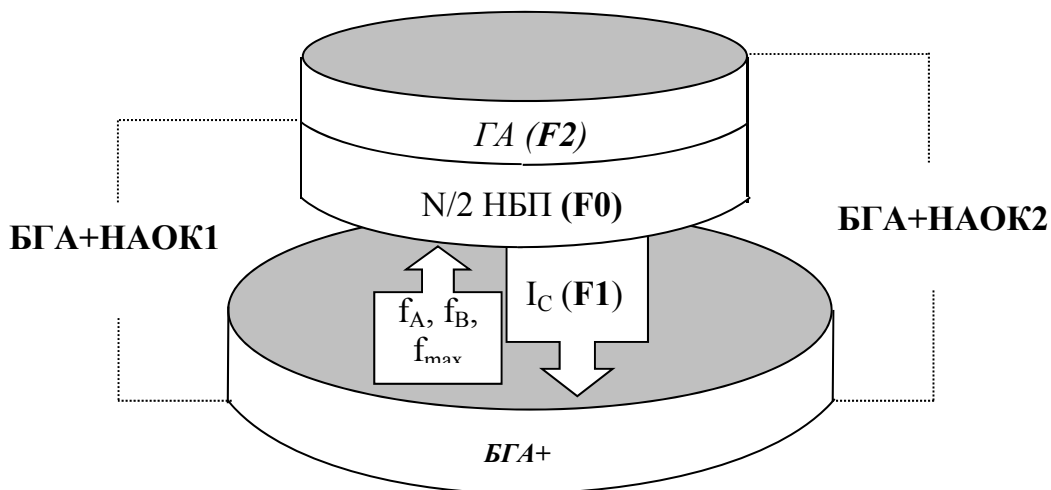


Рис. 8. Двухуровневая модель ГА с нечетким адаптивным кроссовером

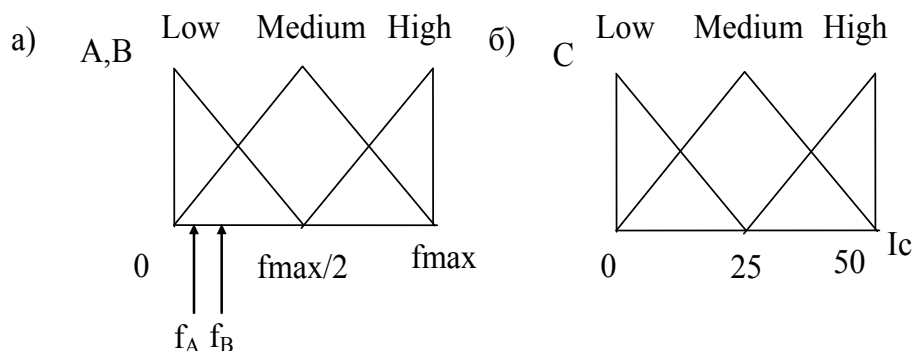


Рис. 9. Входные (а) и выходные (б) лингвистические термы

8.4.3. Последовательный БГА+ с НАОК

На рис. 10 представлен БГА+ (предложенный в п. 8.2) с нечеткими блоками F0, F1, F2. Блоки основных операций обозначены так же как T1-T5.

Генерация нечетких баз правил (F0)

Блок предназначен для первоначальной генерации одной НБП (НАОК 1) и популяции НБП (НАОК 2). Генерация популяции НБП осуществляется произвольным образом в части выходной переменной и жестко задается в части входных переменных, тем самым, исключая возможность появления некорректных НБП вследствие последующего

применения генетических операторов. Фактически, ГА верхнего уровня оперирует и изменяет только консеквентную (выходную) часть правил.

Нечеткий логический вывод (F1)

Под нечетким логическим выводом будем понимать последовательное выполнение трех процессов: фаззификации, нечеткого вывода и дефаззификации, используемых в НАОК для получения значения интенсивности однородного кроссовера по входным показателям эффективности родительской пары индивидов. Нечеткий логический вывод для набора правил, представленного в табл. 4, проиллюстрирован на рис. 11. При нечетком выводе выбранной родительской паре случайным образом назначается одна из НБП популяции. НБП для НАОК содержат набор правил вида «Если А И В, то С». Здесь А и В – антецеденты, С – консеквент.

Фаззификация предназначена для определения степени принадлежности поступающего на вход значения (входной ситуации) элементу входного множества лингвистических переменных. В дальнейшем при рассмотрении используются три стандартные функции степени принадлежности, приведенные на рис. 12.

При задании этих функций помимо типа требуется задание значений отмеченных точек по горизонтальной шкале, по вертикальной шкале значения этих точек 0 или 1.

Нечеткий вывод состоит из ряда последовательно выполняемых процессов. Из всей НБП требуется выбрать правила, которые используют лингвистические переменные, степени принадлежности которых имеют ненулевые значения. Левая часть правила приводится к следующему виду: $\mu_{vx} = \min(\mu_i)$, где i – индекс входной переменной. В случае использования типа импликации $Vx \rightarrow B$: $\min(Vx, Vx')$ результирующая степень принадлежности для одного правила может быть получена следующим образом: $\mu = \min(\mu_{vx}, \mu_{vx'})$.

На следующем этапе среди правил, приводящих к одному решению, должно быть выделено оптимальное. Учитывая наличие в качестве связки решающих правил логической операции ИЛИ, окончательное значение степени принадлежности по одному значению выходной лингвистической переменной может быть получено следующим образом:

$$\mu' = \max_j \min(\mu_{ex}, \mu_{ex'}) ,$$

где j – номер правила.

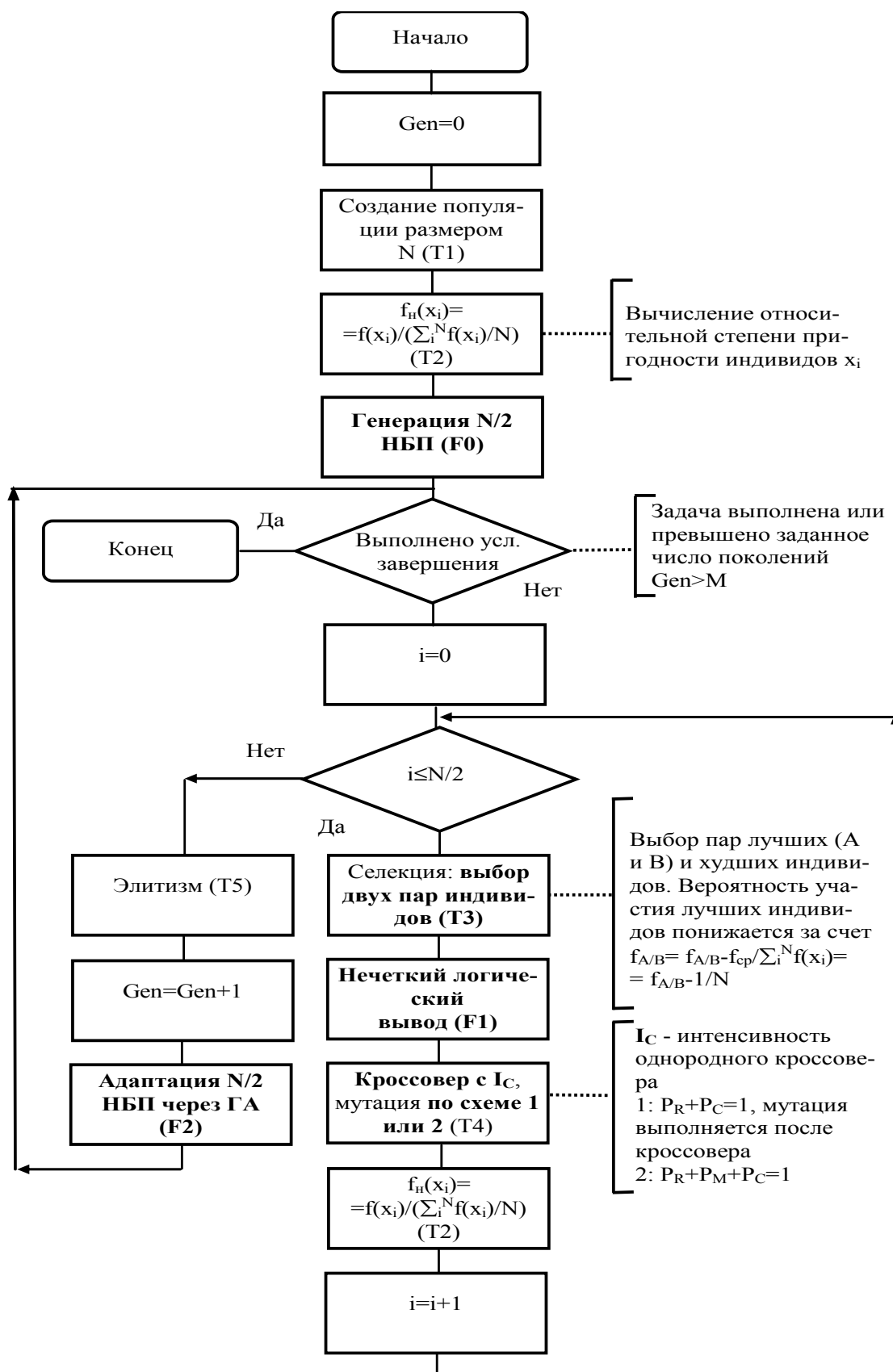


Рис. 10. БГА+НАОК

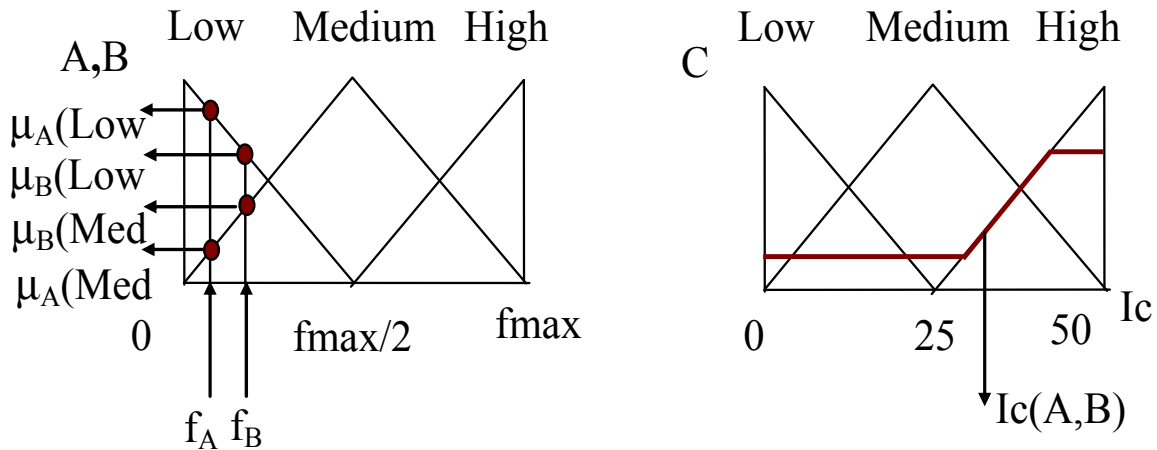


Рис. 11. Нечеткий логический вывод

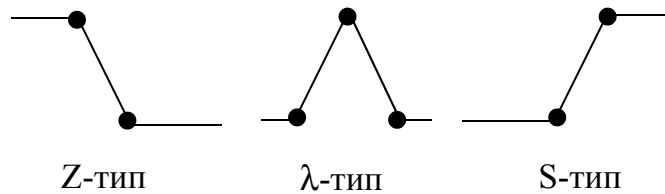


Рис. 12. Используемые функции принадлежности

После перебора активных правил и вывода решения по всем активным правилам строится кривая, которая является композицией всех решений.

Дефаззификация заключается в получении окончательного решения и конкретного значения выходной лингвистической переменной, которое далее будет использоваться для управления объектом. В работе используется один из общепринятых методов дефаззификации – метод центра тяжести.

Адаптация популяции НБП с помощью ГА верхнего уровня (F2).

Адаптация популяции НБП также осуществляется с помощью БГА с модифицированным оператором селекции. От основного ГА работа ГА верхнего уровня отличается следующим:

1. Степень пригодности индивида формируется как сумма степеней пригодностей результатов применения НАОК к родительской паре (наследников).

2. Работа с нечеткими базами правил обуславливает целочисленный способ кроссовера, который фактически сводится только к обмену консеквентных частей, при этом невозможно появление точек скрещи-

вания внутри кода консеквента, а также целочисленный способ мутации, когда невозможна мутация отдельных битов целочисленного представления консеквента. В случае мутации консеквента определен следующий алгоритм: при кодировке консеквента в диапазоне $[0, P]$, если код консеквента равен 0, то результат мутации – 1; если код консеквента равен P , то результат мутации – $P-1$; иначе с вероятностью 50% код консеквента уменьшается или увеличивается на 1.

3. Алгоритм верхнего уровня содержит один шаг (поколение) и выглядит так, как показано на рис. 13.

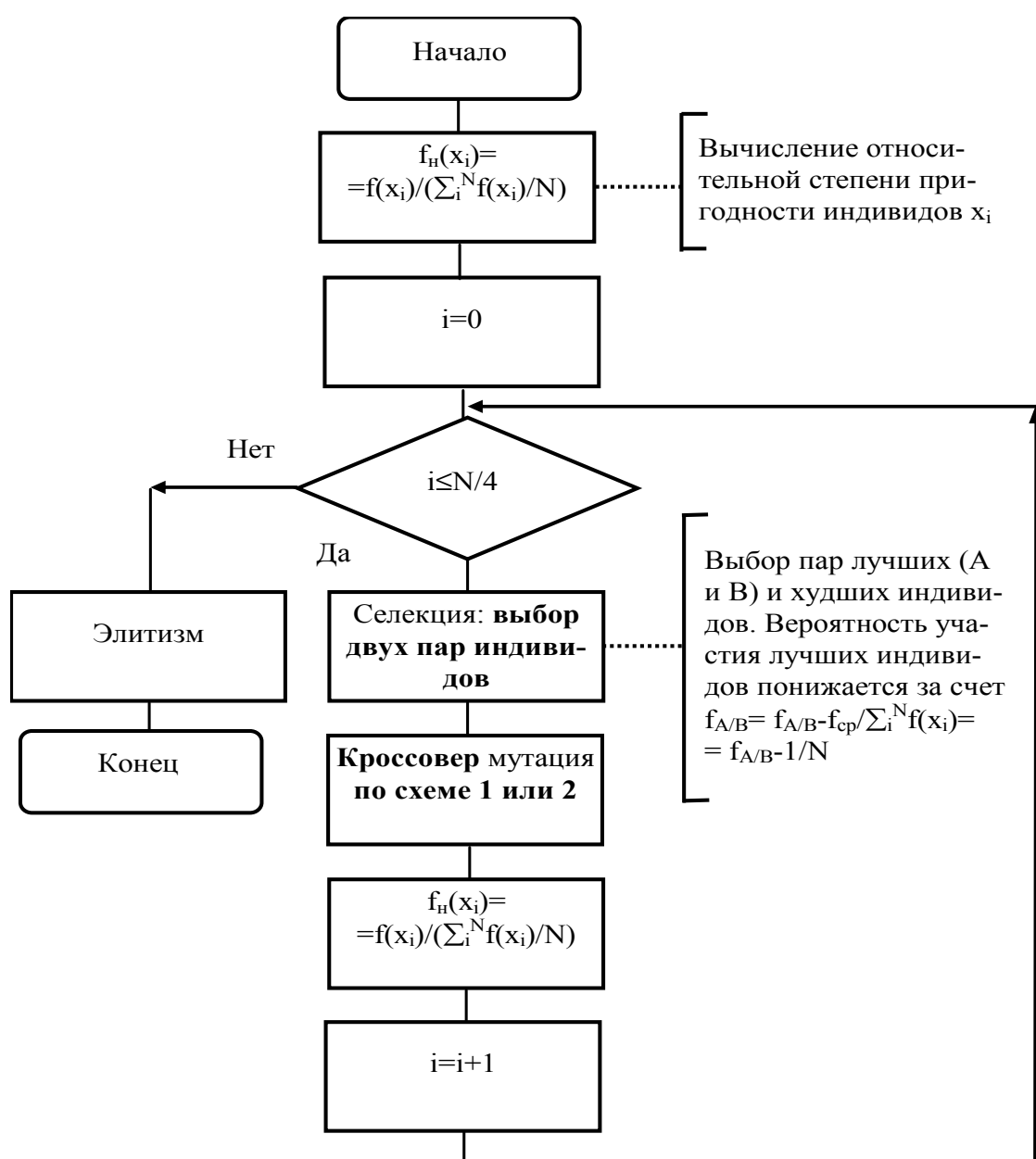


Рис. 13. ГА верхнего уровня (F2)

8.4.4. Результаты экспериментов

Для оценки алгоритмов использовались функции onLine и offLine, определенные в [21]:

$$\begin{aligned} \text{onLine}(T) &= 1/Tn \sum_{t=1}^T \sum_{i=1}^n f(t,i), \\ \text{offLine}(T) &= 1/T \sum_{t=1}^T f(t^*), \end{aligned}$$

где $f(t,i)$ – значение функции степени пригодности индивида i в популяции t , $f(t^*)$ – лучшее значение в поколении t .

Тестирование проходило на тестовой задаче Artificial Ant [1] при параметрах, указанных в п. 8.1.2. с усреднением результатов по 500 запускам, число поколений (число шагов) каждого запуска – 100. В качестве базового принят следующий набор параметров:

- тип селекции – модифицированный;
- тип кроссовера – однородный, вероятность – 0.9;
- вероятность мутации – ≈ 0.01 ;
- тип схемы – 1;
- размер популяции – 60;
- число поколений – 100.

Тест 1. Сравнение эффективности НАОК 1 и НАОК 2 при показателях эффективности индивидов PDM' (рис. 14 а) и PDM'' (рис. 14 б).

Результаты теста 1 позволяют сделать следующие выводы:

1. Использование функции PDM'' позволяет повысить эффективность ГА.

2. В случае использования функции PDM'' эффективность НАОК 1 и НАОК 2 по оценочной функции offLine одинакова, а по оценочной функции offLine НАОК 1 превосходит НАОК 2. Это позволяет сделать выбор в пользу НАОК 1 в случае использования функции PDM''.

3. НАОК 2 доказывает свои адаптивные свойства, показывая устойчивость в значении эффективности по двум оценочным функциям при использовании показателей эффективности PDM' и PDM''.

4. Эффективность НАОК 1 в случае использования PDM'' подтверждает правильность формирования НБП, на основании результатов исследования (см. приложение 2).

Неэффективность НАОК 1 в случае использования PDM' объясняется рассогласованием между динамической шкалой показателей эффективности индивидов и зависимостей по отношению к фиксированному критерию качества решения задачи.

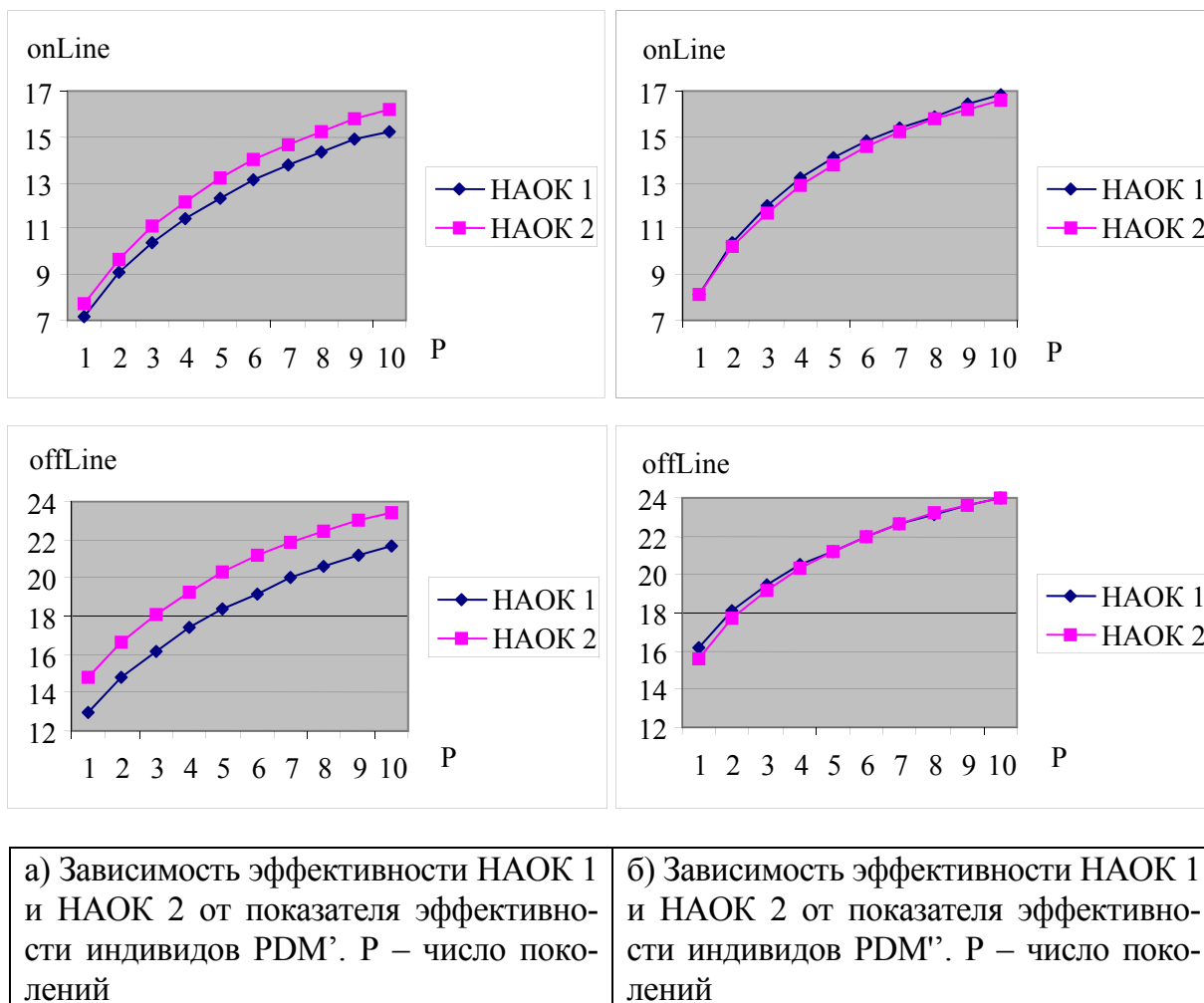


Рис. 14

Тест 2. Эффективность выбора НБП для НАОК 1 (рис. 15). В данном тесте НАОК 1' использует произвольным образом заданную НБП. Эффект настройки нечеткой базы правил выражается в повышении эффективности БГА+НАОК1 по оценочным функциям *onLine* и *offLine*.

Тест 3. Сравнительные оценки эффективности БГА+ и БГА+НАОК 1 (рис. 16). Результаты теста показывают, что БГА+НАОК 1 эффективнее БГА+ по критерию качества решения и при этом имеет большую устойчивость значений функции *onLine*.

Таким образом, БГА+НАОК 1 является предпочтительным по отношению к БГА+НАОК 2 и БГА+ с точки зрения эффективности и предпочтительным по отношению к БГА+НАОК 2 с точки зрения вычислительной сложности алгоритма, поскольку не требует наличия блока F2 и снижает затраты памяти в блоке F0.

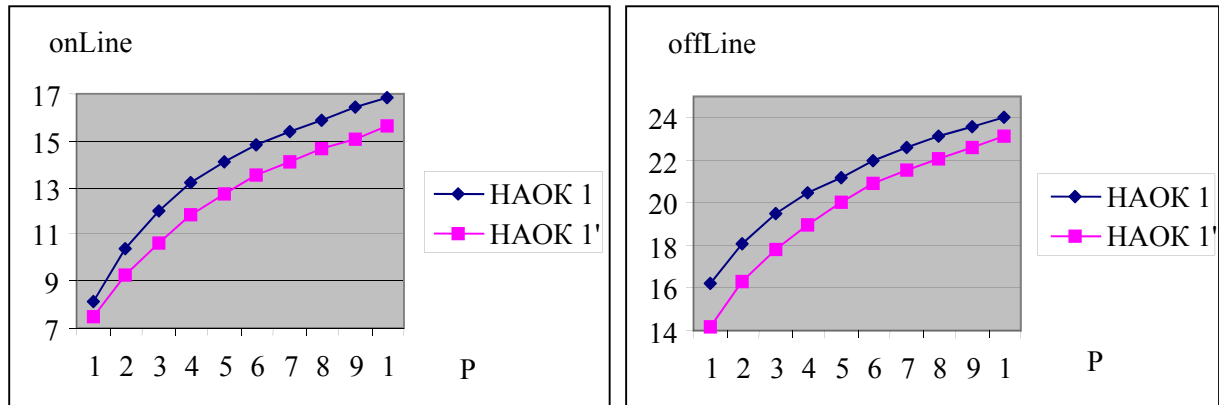


Рис. 15. Зависимость эффективности НАОК1 от выбора НБЗ.
Р – число поколений

Вычислительную сложность ГА представим в количестве последовательных операций. Индекс операции соответствует номеру блока на рис. 11. Тогда вычислительную сложность для разных типов ГА с размером популяции N и числом поколений M можно представить следующим образом:

1. БГА+: $T_{BC'} = T1 * N + T2 * N + M(T5 + N/2 * (T3 + T4 + T2))$
2. БГА+НАОК1: $T_{BC''} = T1 * N + F0 + T2 * N + M * (T5 + N/2 * (F1 + T3 + T4 + T2)) = T_{BC'} + F0 + MN/2 * F1$.
3. БГА+НАОК2: $T_{BC'''} = T1 * N + N/2 * F0 + T2 * N + M(T5 + N/2 * (F1 + T3 + T4 + T2 + F2))$, где $F2 \approx T_{BC'} / (2M)$. Тогда

$$T_{BC'''} = T_{BC'} + N/2 * F0 + MN/2 * F1 + T_{BC'} / (2M) = T_{BC'} * (2M + 1) / 2M + N/2 * F0 + MN/2 * F1$$

На рис. 17 представлены гистограммы примерных зависимостей $T_{BC'}$, $T_{BC''}$, и $T_{BC'''}$ от размера популяции и числа поколений.

Принято допущение, что вычислительная сложность составляющих T1-T5 и F0, F1 равна 1.

Как видно из гистограмм, по вычислительной сложности БГА+НАОК 1 также является более предпочтительным, чем БГА+НАОК 2. Однако он уступает БГА+. Эта проблема решается в

работе за счет перехода от последовательного БГА+НАОК 1 к параллельному БГА+НАОК 1 (ПБГА+НАОК 1).

При этом распараллеливание можно реализовать как по структуре самого ГА, так и в блоке нечеткого логического вывода F1, поскольку именно этот блок увеличивает количество последовательно выполняемых операций на величину $MN/2 \cdot F1$.

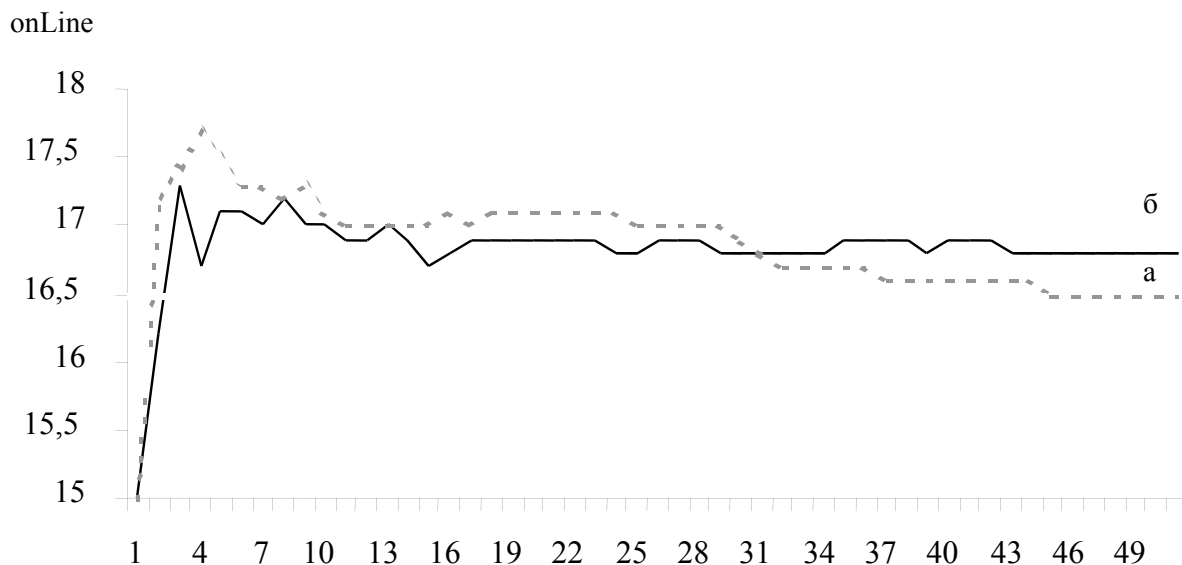


Рис. 16. Зависимость значения функции *onLine* при $P=100$ от числа запусков алгоритмов: а) БГА +; б) БГА+НАОК1

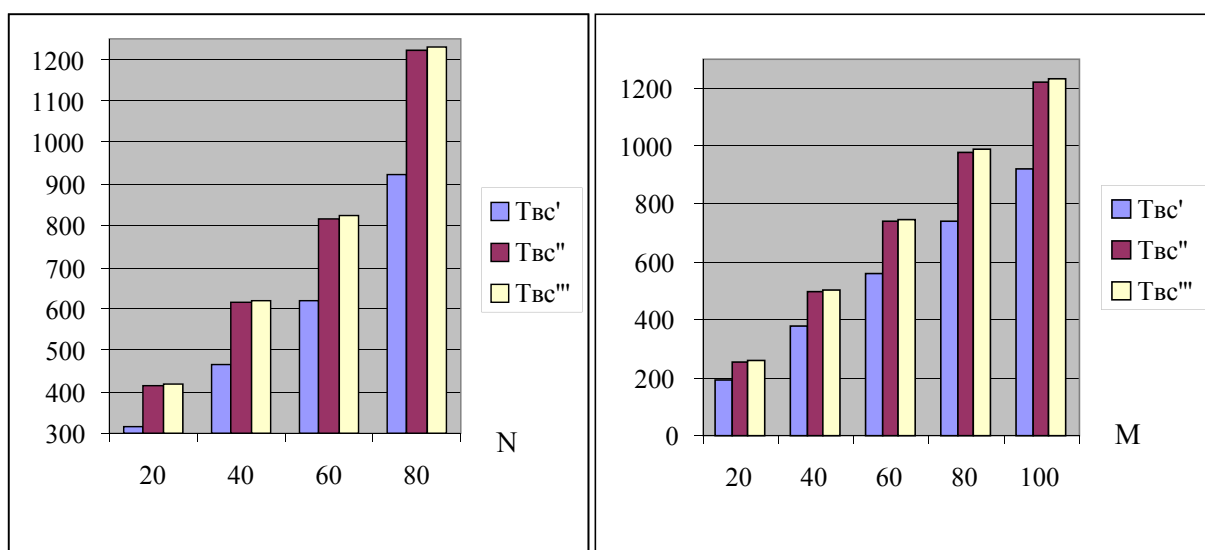


Рис. 17. Гистограммы примерных зависимостей $T_{bc'}$, $T_{bc''}$, и $T_{bc'''}$ от размера популяции (N) и числа поколений (M)

Выводы

1. Основополагающие теоремы определяют влияние генетических операторов селекции, кроссовера и мутации на эффективность ГА. Получено практическое подтверждение теоретического обоснования существенного влияния кроссовера, а также его типа на эффективность ГА. Обоснован выбор однородного кроссовера.

2. Для исследования влияния способа селекции, схемы применения и типов генетических операторов базовый ГА (БГА) преобразован в обобщенный БГА, который отличает наличие двух схем, определяющих последовательность применения генетических операторов.

3. Результатом исследования операторов селекции явилось экспериментальное подтверждение эффективности элитизма, а также разработка обобщенного БГА с модифицированным оператором селекции (БГА+). Его использование позволяет повысить эффективность ГА на 3,6% по оценочной функции *onLine* и на 21,4% по оценочной функции *offLine*. Предлагаемый оператор селекции позволяет избежать формирования промежуточного набора в отличие от БГА, это приводит к экономии памяти для хранения промежуточного набора, а также позволяет сохранить принцип пропорциональности в отличие от SSGA.

4. Результаты исследований обобщенного БГА+ позволяют определить значения параметров, при которых ГА наиболее эффективен с точки зрения качества решения: схема 1 (последовательное применение кроссовера и мутации), вероятность мутации – 0,01, вероятность кроссовера – 0,9, размер популяции – 60, однородный тип кроссовера. Результаты проведенных в работе исследований совпадают с выводами авторов, проводящих аналогичные исследования [26, 27, 29].

5. Результаты тестирования двух разработанных вариантов нечеткого адаптивного однородного кроссовера (НАОК 1 и НАОК 2) позволяют определить следующее:

- при использовании показателей эффективности индивидов популяции PDM' и PDM'' модифицированный базовый ГА с нечетким адаптивным однородным кроссовером второго типа (БГА+НАОК 2) доказывает свои адаптивные свойства, показывая устойчивость по оценочным функциям *onLine* и *offLine*;

- эффект настройки нечеткой базы правил выражается в повышении эффективности модифицированного базового ГА с нечетким адаптивным однородным кроссовером первого типа (БГА+НАОК 1)

на 7,7% по оценочной функции *onLine*, на 3,9% по оценочной функции *offLine*;

– хорошие показатели эффективности НАОК 1 в случае использования PDM'' подтверждают правильность формирования нечеткой базы правил, на основании результатов исследования (см. приложение 2);

– БГА+НАОК 1 более предпочтителен: по качеству решения по отношению к БГА+НАОК 2 и БГА+ и по времени реализации алгоритма по отношению к БГА+ НАОК 2.

6. БГА+НАОК 1 уступает БГА+ по времени реализации. Эта проблема может быть решена за счет перехода от последовательного БГА+НАОК 1 к ПБГА+НАОК 1. Распараллеливание должно быть проведено как по структуре самого ГА, так и в блоке нечеткого логического вывода F1, поскольку этот блок увеличивает количество последовательно выполняемых операций на величину $MN/2 \cdot F1$, где N – размер популяции, M – число поколений ГА.

Список литературы

1. *Koza John R.* Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, 1992. 819 p.

2. *Куприянов М.С., Матвиенко Н.И.* Построение самоорганизующихся систем на основе моделей мягких вычислений // Тезисы докладов научно-технической конференции «Диагностика, информатика, метрология и экология», 1996. С. 171–177.

3. *Куприянов М.С., Матвиенко Н.И.* Использование эволюционных алгоритмов для решения оптимизационных задач // Тезисы докладов научно-технической конференции «Диагностика, информатика, метрология и экология», 1997. С. 183–184.

4. *Muehlenbein H., Schlierkamp-Vooser D.* Analysis of Selection, Mutation and Recombination in Genetic Algorithms. In: Evolution as a Computational Process / Banzhaf W., Eeckman Frank H. (Eds.). Berlin, 1995.

5. *Goldberg D.E.* Genetic Algorithms in Search, Optimization, and Machine Learning. New York: Addison-Wesley, 1989.

6. *Price G.R.* Selection and covariance. Nature, 1970. № 227. Pp. 520–521.

7. *Price G.R.* Extension of covariance selection mathematics. Annals of Human Genetics. 1972. № 35. Pp. 485–489.

8. *Altenberg L.* The evolution of evolvability in genetic programming. In: *Advances in Genetic Programming*. K.E. Kinnear (ed.). Cambridge, MA:MIT Press, 1994. Pp. 47–74.
9. *Altenberg L.* The Schema Theorem and Price's Theorem.
10. *De Jong Kenneth A.* An Analysis of the Behavior of a Class of Genetic Adaptive Systems. Doctoral Thesis. University of Michigan:Dept. Computer and Communication Sciences, 1975.
11. *Eshelman L.J., Caruana R.A., Schaffer J.D.* Biases in the crossover land-scape // *Proc. of the 3rd Intern. Conf. on Genetic Algorithms*. Schaffer J.D. (ed.). Morgan Kaufmann, 1989. Pp. 10–19.
12. *Eiben A.E., Rane P.-E., Ruttkay Zs.* Genetic Algorithms with Multi-parent Recombination // *Proc. of 3rd Intern. Conf. on Parallel Problem solving from Nature (PPSN)*, *Lecture Notes in Computer Science*. Springer Verlag, 1994. Vol. 866. Pp. 78–87.
13. *Jones T.C.* Crossover, macromutation and population-based search // *Proc. of the 6th Intern. Conf. on Genetic Algorithms*. Eshelman L.J. (ed.). 1995.
14. *Cordon O., Herrera F.* A General Study on Genetic Fuzzy Systems // *In Genetic Algorithms in Engineering and Computer Science*. Winter G., Periaux J., Galan M., Cuesta P. (Eds.). Wiley and Sons, 1995. Pp. 33–57.
15. *Куприянов М.С., Матвиенко Н.И.* Оптимальные параметры генетических алгоритмов // *Сб. докл. междунар. конф. по мягким вычислениям (SCM'98)*, С.-Петербург., Июнь 1998. СПб., 1998. С. 283–286.

ГЛАВА 9

ОБЩАЯ КОНЦЕПЦИЯ И ПОДХОД К ПОСТРОЕНИЮ РАСПРЕДЕЛЕННЫХ САМООРГАНИЗУЮЩИХСЯ ИНФОРМАЦИОННЫХ СИСТЕМ (РСИС)

М.С. Куприянов, Ю.А. Шичкина

По определению, интернет вещей (Internet of Things, IoT) – это концепция вычислительной сети физических объектов («вещей»), оснащённых встроенными технологиями для взаимодействия друг с другом или с внешней средой, рассматривающая организацию таких сетей как явление, способное перестроить экономические и общественные процессы, исключаящее из части действий и операций необходимость участия человека.

Сам термин «Интернет вещей» (Internet of Things) был предложен Кевином Эштоном в 1999 году. В этом же году был создан Центр автоматической идентификации (Auto-ID Center), занимающийся радиочастотной идентификацией (RFID) и сенсорными технологиями, благодаря которому эта концепция и получила широкое распространение.

В 2008-2009 произошел переход от «Интернета людей» к «Интернету вещей», т.е. количество подключенных к сети предметов превысило количество людей.

В 2003 году на нашей планете проживало около 6,3 млрд человек, а к Интернету было подключено 500 млн устройств. Разделив количество подключенных устройств на величину населения земного шара, мы увидим, что на каждого человека тогда приходилось по 0,08 такого устройства. Таким образом, в соответствии с определением Cisco IBSG, в 2003 году Интернета вещей еще не было. Смартфоны в то время только появились на рынке.

В 2010 году в результате стремительного распространения смартфонов и планшетных компьютеров количество подключенных устройств выросло до 12,5 млрд, тогда как население Земли составило

6,8 млрд человек. Таким образом, впервые в истории на каждого человека стало приходиться более одного подключенного устройства (1,84 устройства на душу населения).

По оценке Gather, к 2020 году на планете к глобальной сети будут подключены порядка 50 млрд устройств. Для достижения этой цели потребуется развертывание совершенно иной инфраструктуры.

Чтобы получить максимальную отдачу от использования решений Интернета Вещей во всех направлениях человеческой деятельности, необходимо обеспечить современный уровень анализа данных и новый класс интеллектуальных приложений. Эти приложения должны быть не только основаны на новейших методах интеллектуальной обработки данных, они должны быть способны к самоанализу, самоорганизации, самоуправлению.

В данном разделе рассматриваются:

- Общие принципы построения РСИС.
- Модели организации IoT
- Пример задачи, возникающей при построении РСИС в рамках концепции IoT: генерация правил передачи данных в сети Disruption Tolerant Network (DTN) на основе статистической информации

9.1. Традиционные подходы к построению информационных систем

В последние годы все большее значение приобретают задачи управления сложными объектами и процессами в условиях полной или частичной неопределенности, которая может быть априорной или возникать в процессе функционирования системы. Кроме того, современные системы отличаются значительным числом составляющих их элементов и связей между ними, разнообразием состояний системы. К ряду систем дополнительно предъявляется требование живучести – способности системы выполнять установленный минимальный объем своих функций при внешних воздействиях, не предусмотренных условиями нормальной эксплуатации, противостоять таким воздействиям, осуществлять выбор оптимального режима функционирования за счет собственных внутренних ресурсов, перестройки структуры, изменения функций отдельных подсистем и их поведения. Традиционные подходы к решению такого рода задач предполагают использование детерминистического алгоритма управления и применения высокопроизводитель-

ных ЭВМ. Другое направление связано с разработкой более эффективных принципов управления системами, среди которых можно выделить использование принципов самоорганизации. Понятие самоорганизации идет из биологических наук, где она представляет собой основное свойство биологических систем, и охватывает в наиболее общем виде все специфические свойства жизни – сохранение индивидуальности, раздражимость, воспроизведение себе подобных, адаптация к изменяющимся условиям внешней среды, обучение с использованием прошлого опыта. В технических системах термин «самоорганизация» применяется начиная с 60-х годов прошлого века, однако несет в себе различную семантику. Н. Винер [1] определял самоорганизацию как «втягивание в синхронизм», т.е. образование единого ритма работы многочисленных и разрозненных до этого элементов системы, А.Г. Ивахненко [2] сформулировал определение самоорганизации как процесса самопроизвольного увеличения порядка или организации в системе, состоящей из множества элементов, происходящее под влиянием внешней среды. Исходя из анализа литературы, предлагается следующее определение данного понятия в технических системах: самоорганизующаяся система – это система, состоящая из множества связанных между собой элементов, которая самостоятельно под воздействием внешней среды и внутренних микросостояний, на основании накопленного опыта повышает степень своей организации с целью реализации заложенных в ней функциональных возможностей.

Исследования в области самоорганизующихся систем позволили сформулировать следующие основные принципы их построения:

1. Свойство необходимого разнообразия состояний элементов системы, согласно которому самоорганизующаяся система должна включать в себя большое множество элементов, необходимое разнообразие входящих элементов, наличие большого количества прямых и обратных связей между элементами, наличие элементарных функций у отдельных элементов.

2. Непрерывная приспособляемость системы к изменяющимся внутренним и внешним условиям существования и непрерывное совершенствование поведения при неизменных условиях с учетом предыдущего опыта.

3. Отбор элементов как на уровне выбора элементов для образования системы (селекционная самоорганизация), так и на уровне выбора элементов из существующих для решения конкретных задач (элекционная самоорганизация).

4. Наличие у самоорганизующейся системы цели функционирования и ресурса для ее достижения (целенаправленное поведение системы).

5. Обособленность системы от окружающей среды и ее целостность.

6. Иерархическое строение.

9.2. Классификация самоорганизующихся систем

Основываясь на свойствах описанных в литературе систем, все самоорганизующиеся системы по степени самоорганизации можно классифицировать на следующие четыре основные группы:

1. Адаптивные системы – система, которая сохраняет работоспособность в условиях непредвиденного изменения свойств управляемого объекта, цели управления и условий окружающей среды посредством смены алгоритмов своего функционирования или поиска оптимальных состояний [3].

2. Обучающиеся системы – система в которой информация о неизвестных характеристиках процесса или окружающей обстановке появляется в ходе обучения, а полученный опыт используется в дальнейшем для оценки, классификации, принятия решений или управления так, что качество работы системы улучшается.

3. Самообучающиеся системы – система, функция которой вырабатывается и совершенствуется в процессе самообучения. Отличается от обучающейся отсутствием внешней корректировки и дополнительной информации о верности реакции системы.

4. Системы с искусственным интеллектом – это системы, проявляющие такие интеллектуальные способности, как принятие сложных решений, планирование и другие. Такие системы не только вырабатывают оптимальный алгоритм, но и выбирают стратегию обучения в соответствии с заложенной в них целевой функцией [4 -16].

9.3. Математическая модель самоорганизующейся информационной системы (СИС)

Математическую модель СИС можно представить в виде восьмерки:

$$S = \{\Sigma, F, \Theta, \Omega, A, P, T, \Psi\},$$

где Σ – структура системы, F – функция системы, Θ – отношение эмерджентности, Ω – множество целей системы, A – отношение адаптивности, P – множество элементов памяти, T – множество моментов времени, Ψ – множество ЛКТ.

Понятие структуры Σ отражает упорядоченность, организованность системы. Под структурой системы будем понимать совокупность элементов и связей между ними, которые определяются, исходя из распределения функций и целей, поставленных перед системой. Обычно структура системы представляется графически в виде графа с множеством вершин, соответствующих элементам системы, и множеством ребер (дуг), соответствующим связям между элементами.

Для структуры СИС характерны следующие свойства: целостность, адаптивность и развитие. Целостность структуры определяется множеством элементов и разветвленностью связей между ними, и проявляется двумя основными закономерностями: целое не сводится к простой сумме частей и изменение в одной части вызывает изменение во всех остальных частях и во всей системе.

Свойство адаптивности структуры обуславливает внутренние преобразования структуры, которые реализуются ее элементами, не выходя за пределы ее границ и направлены на сохранение и улучшение ее функционирования.

Под развитием структуры понимается усложнение системы, накопление ею информации, переход в более упорядоченное состояние.

Функция системы F является, в отличие от структуры, внешним проявлением свойств системы при взаимодействии с объемлющей средой. Функция системы представляет собой способ действия системы при достижении цели. Функция СИС обуславливается ее целью. Функции, реализуемые СИС могут быть условно разделены на три группы: целевая, основные и вспомогательные. Целевая функция системы соответствует основному функциональному назначению системы и задается исключительно целью функционирования системы. Основные функции отражают ориентацию системы и представляют собой совокупность макрофункций, реализуемых системой и необходимых для наиболее оптимального достижения цели. Основные функции системы определяются целевой функцией и критериями качества системы. Под дополнительными функциями будем понимать функции, направленные на поддержание качества функционирования системы в некоторых пределах или его повышение.

Отношение эмерджентности Θ отражает единство структуры и функции системы, связь между ними; из структуры и функции образует систему. Эмерджентность некоторым образом соединяет элементы структуры с функциями, отображает F в Σ :

$$\Theta \subset |F| \times |\Sigma|.$$

Множество целей Ω соответствует множеству целей, стоящих перед системой, и может рассматриваться как дистрибутивная решетка, наделенная операциями объединения, пересечения и композиции.

Адаптивность A связывает поведение и структуру системы с изменениями воздействий внешней среды и с внутренними состояниями. Адаптивность некоторым образом отображает воздействия внешней среды X и состояние M в структуру Σ и функцию F . Отношение A распадается на два: A_1 и A_2 , где $A = A_1 \times A_2$:

$$\begin{aligned} A_1 &= |\chi| \times |\Sigma|; \\ A_2 &= |\chi| \times |\Sigma|, \end{aligned}$$

где $\chi = S_{\text{вх}} \times M$, $X \subseteq \chi$.

Время T – направленное множество моментов времени, в котором каждому моменту поставлена в соответствие определенная ситуация.

Память P . Для использования предыдущего опыта система должна иметь возможность для накопления и хранения информации, т.е. обладать памятью. С математической точки зрения память системы может быть рассмотрена как дистрибутивная решетка с операциями объединения и пересечения.

Множество ЛКТ Ψ включает в себя все ЛКТ всех процессов, составляющих систему с привязкой к моментам взятия этих точек (множеству T).

9.4. Структурно-функциональная модель СИС

С учетом вышеизложенного представления математической модели СИС предлагается следующая структурно-функциональная модель СИС, изображенная на рис. 1.

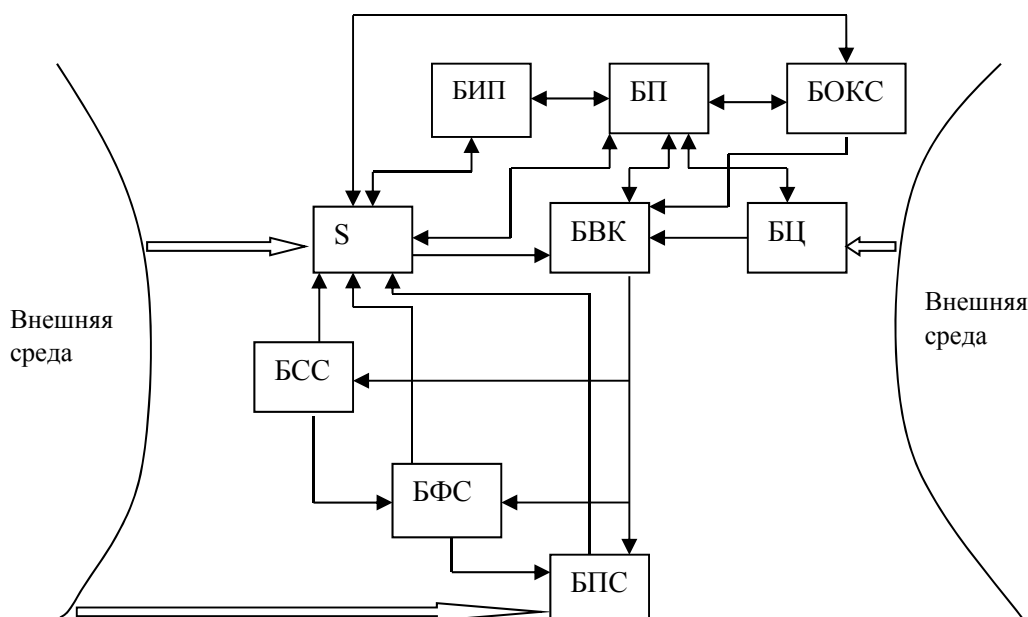


Рис. 1. Структурно-функциональная модель самоорганизующейся системы.

На рисунке 1 обозначено: *S* – самоорганизующаяся система, БИП – блок исходной программы, БСС – блок структурной самоорганизации, БФС – блок функциональной самоорганизации, БПС – блок параметрической самоорганизации, БЦ – блок целей, БВК – блок вычисления критериев, БП – блок памяти, БОКС – блок определения консистентных состояний.

Рассмотрим кратко основные функции каждого из блоков.

Внешняя среда. Под внешней средой будем понимать множество элементов, которые не входят в систему, но изменение их состояния вызывает изменение состояния системы.

Самоорганизующаяся система *S*. На ней при помощи остальных блоков системы строится модель объекта или процесса, реализуются эвристические процессы и алгоритмы самоорганизацию. Система состоит из множества элементов, определяемого сложностью задачи, с изменяющейся функцией элемента и множеством регулируемых связей. Параметры связей и элементов, их функции изменяются в зависимости от класса задач, конкретных условий задачи и условий внешней среды при решении задачи. Перестройка системы осуществляется в зависимости от состояния других блоков системы.

Блок исходной программы БИП. Блок БИП задает предварительную структуру системы, предварительный набор функций эле-

ментов, которые будут применяться при самоорганизации, предварительный набор параметров. Кроме того, он задает ограничения на структуру программы, типы функций элементов, параметры.

Блок структурной самоорганизации БСС. Блок БСС изменяет структуру системы увеличением или уменьшением числа элементов в системе, изменением топологии, количества и характера связей.

Блок функциональной самоорганизации БФС. Блок БФС производит адаптацию типа функций элементов системы согласно заложенным алгоритмам адаптации. Кроме того, он управляет началом процесса структурной самоорганизации при исчерпании ресурсов функциональной самоорганизации и недостижении цели.

Блок параметрической самоорганизации БПС. Блок БПС производит подстройку параметров системы по специальным алгоритмам. При недостижении цели в течение определенного времени он запускает блок функциональной самоорганизации.

Блок целей БЦ управляет выработкой критериев СИС. Цели системы в данном случае задаются извне и являются для системы predetermined.

Блок БВК осуществляет выбор и вычисление критериев, которые определяются целью и выбираются из множества допустимых критериев, которые являются эвристическими и закладываются в БВК при проектировании.

Блок памяти БП осуществляет накопление и хранение информации о прошлом поведении системы, о ситуациях, приведших к достижению цели, также в нем запоминаются необходимые сообщения и ЛКТ.

Блок определения консистентных состояний БОКС осуществляет, в зависимости от используемого механизма взятия локальных контрольных точек (ЛКТ), определение консистентного состояния системы в каждый момент времени и позволяет в случае возникновения отказа или сбоя в системе осуществить возврат к одному из прошлых согласованных состояний [11].

Для того, чтобы система могла отвечать задачам функционирования самоорганизующихся систем, она должна обладать определенными свойствами. К основным свойствам СОС можно отнести:

1. Высокая степень сложности.
2. Функциональная избыточность.
3. Переменность структуры системы.
4. Структурная однородность.

5. Параллельность.
6. Многофункциональность.
7. Иерархичность.
8. Распределенность.

Алгоритм самоорганизации системы (рис. 2) работает следующим образом. В первый момент времени вводится набор сигналов из внешней среды. При этом исходная система формирует некоторый набор значений выходных сигналов. Затем производится обращение к блоку критериев и оценка соответствия набора значений выходных сигналов критериям выходных сигналов системы. В случае, если набор выходных сигналов удовлетворяет набору критериев, выводятся данные и останавливается самоорганизация системы. Но, естественно, маловероятно, что случайно организованная система адекватно решит поставленную задачу с первого обращения к ней. Поэтому, в случае отрицательной оценки близости выходных сигналов к критериям, производится подстройка параметров элементов системы. Число подстроек заносится в счетчик числа подстроек. Если число подстроек не превышает заданного предельного значения, то в память заносятся новые параметры исходной системы. Затем опять оценивается функционирование системы и ее элементов и, в случае необходимости, производится новая подстройка.

Так, циклически осуществляется подстройка параметров, пока число их не превысит ϵ_n . После этого изменяются функции элементов. Динамика подстроек функций организована аналогично подстройкам параметров с тем отличием, что цикл подстроек параметров производится после каждого изменения функции. Цикл изменения функций заканчивается после того, как число их изменений превысит некоторое значение ξ_n . При этом включается в работу последний цикл – цикл изменений структуры системы. Причем на каждое изменение структуры системы приходится цикл изменений функций элементов [12].

В случае если на любом этапе подстроек выходные параметры удовлетворяют критериям, производится вывод данных, запоминание конечной структуры системы и этапов реализации процесса самоорганизации. Если и число изменений структуры системы превысит заданный предел τ_n , а выходные сигналы системы не будут удовлетворять критериям, система выдает запрос на изменение цели функционирования, т.е. показывает, что цель была определена неверно, и останавливается. Общая концепция и подход к построению моделей СИС апробирован в различных областях. Некоторые результаты изложены, в частности, в [10–15].

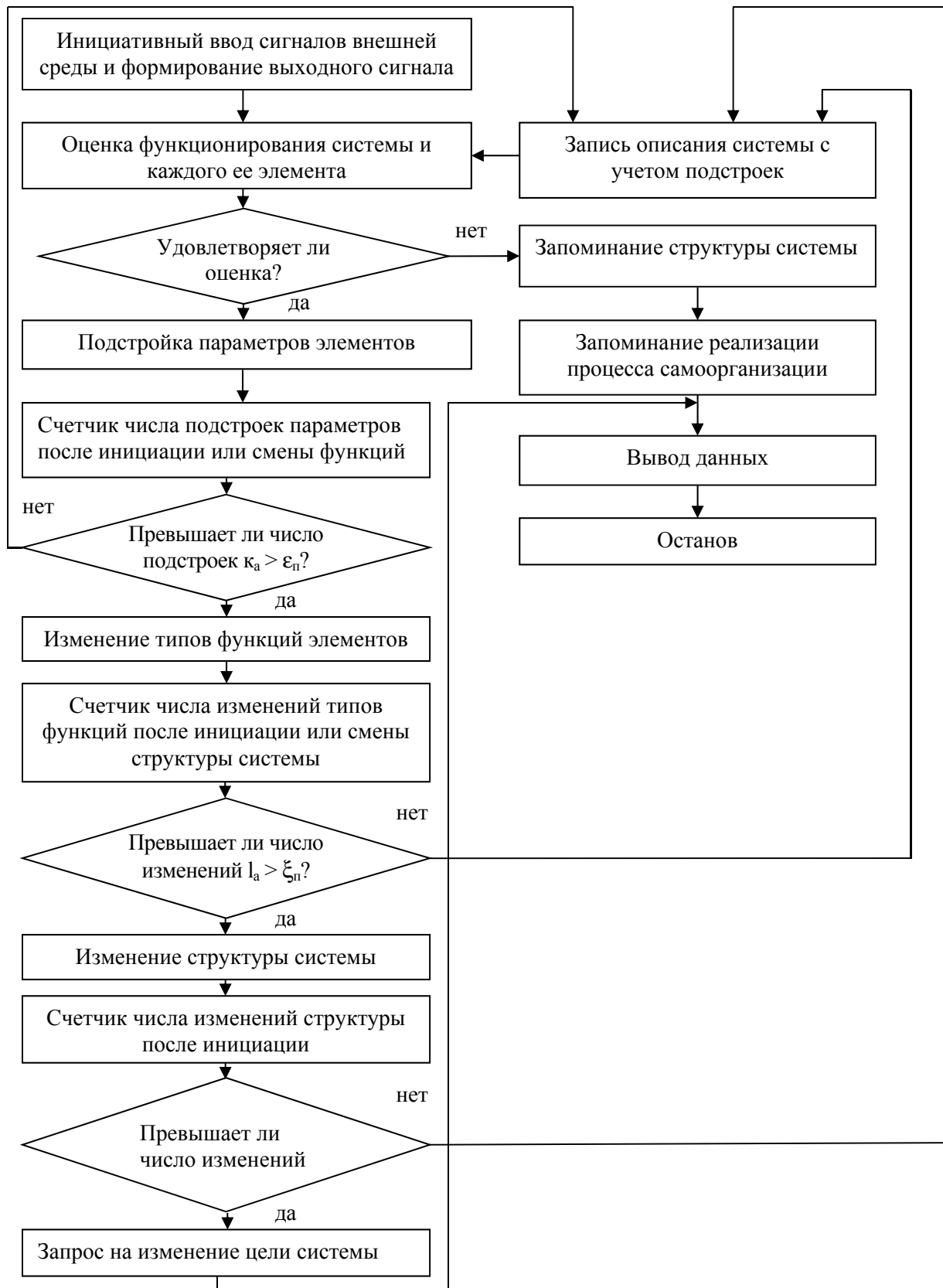


Рис. 2. Алгоритм самоорганизации системы

9.5. Пример построения самоорганизующейся системы обмена информацией

Ускоренный и массовый переход человечества к IoT не может быть осуществлен без адаптации технологий построения информационных систем под концепцию IoT. Информационные системы сегодня охватывают на аппаратном уровне самые разные устройства от смартфонов до кластеров и суперкомпьютеров, несколько разновидностей сетей и протоколов на всех уровнях передачи информации, на программном уровне требуют корректной работы одновременно с различными операционными системами и прикладным программным обеспечением. Поэтому одной из актуальных задач сегодня является построение системы обмена информацией, которая бы позволяла осуществлять не только передачу данных между устройствами РСИС, но и распределение задач между доступными узлами. При этом в свете наличия в РСИС не только стационарных сетей, но и динамических сетей, система обмена информацией должна быть самоорганизующейся, т.е. в зависимости от наличия, состояния сетей и предистории своего функционирования, система должна вбирать наилучший способ обмена информацией.

Рассмотрим модель такой системы на примере задачи генерации правил передачи данных в динамических сетях DTN.

Аналитики предсказывают, что к 2020 году трафик в беспроводных сетях намного превзойдет трафик в проводном Ethernet. Технология WiFi одна из самых перспективных на сегодняшний день в области компьютерной связи. WiFi – это возможность передачи данных между устройствами на короткие дистанции без помощи проводов. Устройства подключенные по беспроводной технологии образуют сеть. Одной из характеристик качества работы такой сети – это технология организации данных по сети, маршрутизация.

В существующих сетях передачи данных, как правило, используется ряд традиционных узкоспециализированных алгоритмов маршрутизации. Основной особенностью работы данных алгоритмов является оценка стоимости маршрута на основе набора метрик, используемых в протоколе.

Задача обеспечения требуемого качества обслуживания в сетях передачи данных решается путем введения правил обработки потоков передаваемой информации в соответствии с определенным типом трафика.

Однако задание статических правил для определенных потоков данных в современных конвергентных сетях недостаточно эффективно [17, 18].

Для решения данной проблемы исследователями Стэнфордского университета разработана концепция программно-конфигурируемых сетей. Она основывается на использовании технологии OpenFlow, которая позволяет осуществлять более гибкое управление сетью за счет анализа потоков данных, проходящих через коммутаторы. На специализированных контроллерах OpenFlow задаются динамические правила, осуществляющие перераспределение потоков данных на менее загруженные маршруты согласно заданным условиям. Данная архитектура также позволяет собирать информацию о состоянии сети в критически важных узлах [19, 20].

Предлагаемые разработчиками OpenFlow решения используют подходы, заложенные в традиционные протоколы маршрутизации, адаптируя и динамически изменяя таблицы правил для составления новых маршрутов. Такой подход учитывает конвергентность сети, но по-прежнему не позволяет классифицировать потоки передаваемых данных в соответствии с заданными критериями (тип приложения или сервиса, объем передаваемых данных, пропускная способность канала, требуемая полоса пропускания, задержки, джиттер и т.д.), а также не дает возможность осуществлять балансировку нагрузки в каналах связи и между сетевыми устройствами. Кроме перечисленных проблем современные алгоритмы маршрутизации должны решать задачи повышения надежности и производительности критически важных сегментов сети при передаче информации реального времени, например потоки данных видео или VOIP-сервисов и другие.

Но, одно дело, когда устройства статические, другое дело, когда устройства 80% своего времени находятся в движении. В этом случае проблем становится еще больше.

Статические алгоритмы маршрутизации, в отличие от динамических, не учитывают постоянно изменяющуюся топологию сети. Это делает ее непригодной для использования в большинстве сетей.

Все алгоритмы используют одну из трех математических моделей – Дейкстра, Беллмана-Форда, Флойда-Уоршелла. Но если статические алгоритмы распространяют их на всю описываемую подсеть, то динамические только локально, используя развитые метрики оптимальности.

Алгоритм заливки является самым надежным и быстрым [21] из всех существующих алгоритмов. Принцип функционирования заклю-

чается в рассылке пришедшего пакета во все линии, кроме той, по которой он пришел. Но его единственный и главный минус – недопустимо большое значение трафика. Данный алгоритм является оценочным при тестировании новых разработок и все ещё используется в специализированных сетях (например, военных).

Алгоритм маршрутизации на основании потока основывается на предположении о том, что трафик внутри сети можно описать неким статистическим законом, на основании которого и выбираются оптимальные схемы маршрута [21].

Одним из наиболее совершенных на сегодняшний момент алгоритмов в маршрутизации является маршрутизация с учетом состояния линий. Метрикой для данного алгоритма является средняя величина задержки для тестового пакета, что отражает не только длину маршрута, но и загрузку канала. Практической реализацией данного алгоритма является протокол OSPF.

В данном разделе рассматривается случай построения динамической сети DTN между устройствами, находящимися в движении и связывающихся между собой посредством Wi-Fi. При этом вместо алгоритма заливки строится система маршрутизации на основе информации о предыдущих передвижениях устройств с применением теории графов.

9.5.1. Постановка задачи генерации правил в динамических сетях

На транспортных средствах, передвигающихся в заданной области установлены устройства, которые связываются между собой посредством Wi-Fi. Данные о передвижении транспортных средств берутся из базы данных Glonass. Изначальный формат csv, пример записи приведен на рис. 3.

Наименования полей:

rxDevice, fileId, txDevice, gentime, txRandom, msgCount, dsecond, latitude, longitude, elevation, speed, heading, ax, ay, az, yawrate, pathCount, radiusOfCurve, confidence

Пример записи:

1566, 1158354, 1566, 291864241884364, -9018, 19, 35800, 42.41119, -83.874352, 269.10001, 18.280001, 176.1125, 9.9999998E-3, 0.56999999, 0.0, 1.86, 9.0, 1829.0, 95

Рис. 3. Наименование полей и пример записи в начальной базе данных

Основная задача заключается в генерацию правил передачи данных в сети DTN от произвольного устройства А к произвольному устройству В посредством Wi-Fi с радиусом действия 100 метров.

Будем полагать, что передача данных должна осуществляться только по каналам с 80% вероятностью доставки. С учетом того, что сеть строится между движущимися устройствами и от этого теряется 100% вероятность доставки по одному выбранному маршруту, а алгоритм заливки исключен по причине перегрузки сети, то строится должно несколько запасных путей.

Основная задача разбивается на ряд следующих подзадач:

1) предварительная подготовка данных, которая включает: подбор модели базы данных для ускоренной обработки данных; перевод данных из формата cvs в формат выбранной базы данных; очистку данных и другие операции над данными, способствующие более быстрому и качественному решению исходной задачи;

2) построение маршрутов с применением аппарата нечеткой логики и теории графов.

9.5.2. Предварительная подготовка данных

Выбор базы данных

Наиболее распространённые SQL базы данных являются нежелательными для данной задачи, так как с увеличением объема информации их производительность заметно падает.

Основным преимуществом баз данных NoSQL является их массовая способность масштабирования. Базы данных NoSQL предлагают больше места для хранения, легко справляются с огромными объемами данных, и предлагают предсказуемую производительность (в 1.5–2 раза большую, в сравнении с SQL).

Для решения поставленной задачи была выбрана NoSQL база данных MongoDB 3.2 под операционной системой Ubuntu Server 14.04 LTS.

Импорт данных в базу данных

Согласно предложенной документации по программе Safety Pilot Model Deployment (SPMD) для набора данных BsmP1 выделены следующие поля данных и применена команда mongoimport:

```
mongoimport --db moto --collection traffic --fields rxDevice,fileId,txDevice,gentime,txRandom,msgCount,dsecond,longitude,latitude,elevation,speed,heading,ax,ay,az,yawrate,pathCount,radiusOfCurve,confidence --type csv --numInsertionWorkers 40 --file april.csv
```

В результате этой операции над файлами тестовой базы данных за апрель и октябрь 2014 г. было импортировано в базу данных MongoDB 2 267 163 140 записей.

9.5.3. Методика извлечения знаний из баз данных. Концепция нанесения сетки

Методика извлечения знаний из баз данных (Knowledge Discovery in Databases (KDD)) берет свое начало во второй половине 20 века и описывает не конкретные алгоритмы, а определенную последовательность действий, которой необходимо следовать для извлечения полезного знания.

Изначально исходные данные необходимо получить из источников данных. Зачастую такие данные необходимо не просто собрать, а консолидировать из нескольких источников. После того, как данные получены и помещены в некоторое хранилище их необходимо очистить от «мусорных» данных (не имеющих ценности для получения полезного знания). На третьем этапе очищенные данные необходимо подвергнуть трансформации. Трансформация представляет собой подготовку очищенных данных к дальнейшей аналитике, поскольку многие из методов предполагают определенный формат данных для их применения. Четвертый этап – это Data Mining, процесс извлечения полезных знаний из исходных данных. На последнем этапе осуществляется интерпретация полученных знаний.

Когда речь идет о больших данных подход KDD можно применить с некоторыми модификациями, а именно пройти этапы очистки и трансформирования данных неоднократно. Данный подход позволит сократить количество полезных данных, что позволит более быстро и эффективно использовать методы извлечения знаний.

Очистка данных

Полученный в BsmP1 набор данных избыточен для решения поставленной задачи. Набор данных содержит 13 типов данных, отражающих различные атрибуты транспортных средств, с помощью которых были получены эти данные.

В файле BsmP1 описаны следующие 19 атрибутов: RxDevice, TxDevice, Gentime, TxRandom, MsgCount, Dsecond, Latitude, Longitude, Elevation, Speed, Heading, Ax, Ay, Az, Yawrate, PathCount, RadiusOfCurve, Confidence.

Из всего набора данных для построения маршрутов необходима только информация об идентификаторах устройств (TxDevice), гео-

данных (Longitude и Latitude) и времени (Gentime). Таким образом из 19 изначальных полей после очистки останется только 3 поля. Из них поле Gentime должно быть преобразовано в два других поля: временная зона суток (TimeZone) и день недели (WeekDay). Получается, что из 19 полей получится в новой базе данных 5 полей, которые будут содержать «полезные» данные.

Также из обработки необходимо исключить некорректные записи, например, записи с координатами, соответствующими географическим полюсам или находящимся в значительном удалении от плотно расположенных данных. Такие записи могли быть получены при некорректной работе устройств геопозиционирования.

Концепция сетки

Главной идеей упрощения обработки данных и ускорения процесса обращения к данным является группирование данных по общему признаку. Поскольку для набора данных BsmP1 общими параметрами, не зависящими от других, являются координаты и время, то будет крайне логичным объединять данные в группы по времени и местоположению.

Самым простым и в то же время правильным будет объединение записей в группы по некоторому участку пространства, в котором устройства будут находиться, например, разбиение плоскости на прямоугольники или квадраты, ввиду двумерности получаемых от GPS-спутника координат. Такое разбиение будем называть далее координатной сеткой (рис. 4).

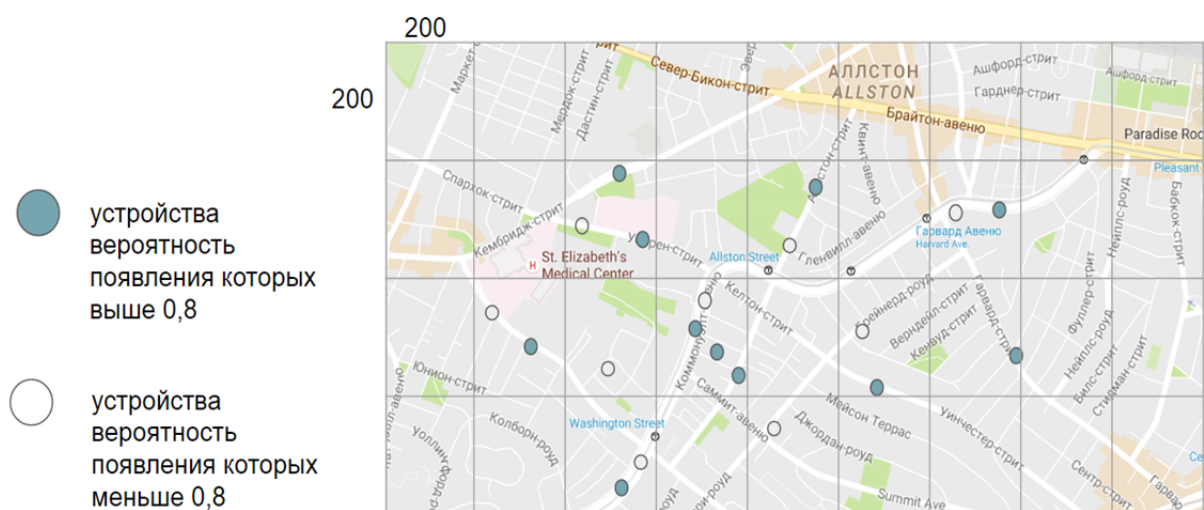


Рис. 4. Вид координатной сетки

Изначально построение маршрута строится между ячейками координатной сеткой. Таким образом мы получаем маршрут, который будет связывать множества устройств, локализованных в соответствующих ячейках координатной сетки. Такой маршрут будем далее называть укрупненным маршрутом. По построенному укрупненному маршруту второй итерацией будут построены уже несколько ранжированных по убыванию вероятности доставки информации конечных маршрутов передачи информации. Такой подход позволит ускорить процесс обработки данных за счет уменьшения объема записей в базе данных при построении укрупнённого маршрута.

Для нанесения координатной сетки на карту местности необходимо выбрать единичную меру, например, метр, и было бы все гораздо проще, если бы градусы долготы и широты однозначно переводились в метрическую систему. Как известно, длина меридиана постоянна, и, исходя из этого, можно однозначно перевести требуемую единицу в градусы широты. Для долготы сложнее, так как длина 1 градуса в метрах изменяется в зависимости от координаты.

Если при нанесении двумерной координатной сетки выполнять точные расчеты, то для обработки каждой записи необходимо заранее произвести расчеты длины метра в зависимости от координат объекта, и хранить все эти данные в отдельном хранилище. Данный подход требует дополнительных хранилищ и обращений к этим хранилищам. Поэтому, для локализованных вычислений координатной сетки в рамках одного города или района, допустимо использовать приближенное среднее значение, исходя из разности крайних координат, деленной на два.

Шаг сетки можно выбрать в зависимости от конечного назначения полученной выборки.

Для ускоренной навигации по координатной сетке в базу данных следует добавить индексируемое поле `nSquare`, которое является уникальным идентификатором ячейки координатной сетки.

Переходя к группировке по времени следует заметить, что загруженность определенных областей различными транспортными средствами, по которым возможно собрать телеметрию геолокации, достаточно периодична относительно дней недели.

Исходя из периодичности данных имеет смысл разбить весь набор данных на 7 частей по дням недели и работать с каждым новым набором в отдельности. Данный подход позволит в дальнейшем рас-

пределить обработку вновь поступающих данных на кластер из нескольких машин, каждая из которых будет обрабатывать меньшую часть данных, например, только один из дней недели.

Также имеет смысл ввести еще одну временную координату помимо дня недели – временная зона суток. Выбор подобной координаты обусловлен не только вопросом упрощения навигации по временной оси, но и для сглаживания погрешностей местонахождения устройства в определенной ячейке координатной сетки.

Таким образом, после трансформирования данных, будет получена новая структура данных, в которой поле `loc` содержит координаты устройства, поле `fileId` содержит идентификатор устройства `RxDevice`, `dayTime` содержит количество полных минут, прошедших с начала суток, `weekDay` – порядковый номер дня недели, `timeZone` – номер временной зоны суток, `nSquare` – уникальный идентификатор сектора координатной сетки и поле `square` содержит координаты сектора в сетке:

После первой трансформации количество записей не уменьшилось, но количество полей изменилось. В результате в новую базу данных попали только необходимые и полезные данные. Данный набор данных назовем предварительной сеткой (`preMesh`).

В дальнейшем над этим набором данных будут произведены действия для глубокой очистки данных методом применения нечетких срезов (`fuzzy sets`), что позволит значительно уменьшить количество данных для ускоренного применения методов извлечения знаний.

Также подобный набор данных в нашем случае обработки данных всего за два месяца позволил распараллелить задачу построения маршрутов на 7 вычислительных узлов, по одной для каждого дня недели.

9.5.4. Визуализация полученных данных

Для оценки полученных данных и принятия решения о дальнейшей целесообразности применения методов извлечения знаний была проведена работа по получению графического представления данных. Этот процесс был реализован с помощью программного модуля, написанного на языке PHP 5.6. На входе в функцию подавалось название коллекции в MongoDB и полученное максимальное значение устройств или записей в ячейках сетки.

Далее приведен пример полученного изображения предварительной сетки за понедельник (рис. 5). В ходе анализа было определено,

что в центре, где расположены наиболее загруженные ячейки сетки, расположен такой крупный город, как Индианаполис.

Получаемые файлы ограничены крайними координатами сетки для конкретного набора. Каждая точка изображения – это отдельный сектор нанесенной на координатной сетке. Благодаря полученным изображениям мы можем понять, где и на сколько много у нас расположено записей, по каким направлениям мы можем рассчитывать на доставку сообщений.

9.5.5. Применение нечеткой логики для прореживания данных в сетке

Большая часть данных, обрабатываемых в современных информационных системах, носят четкий, числовой характер. Однако в запросах к базам данных, которые пытается формулировать человек, часто присутствуют неточности и неопределенности.

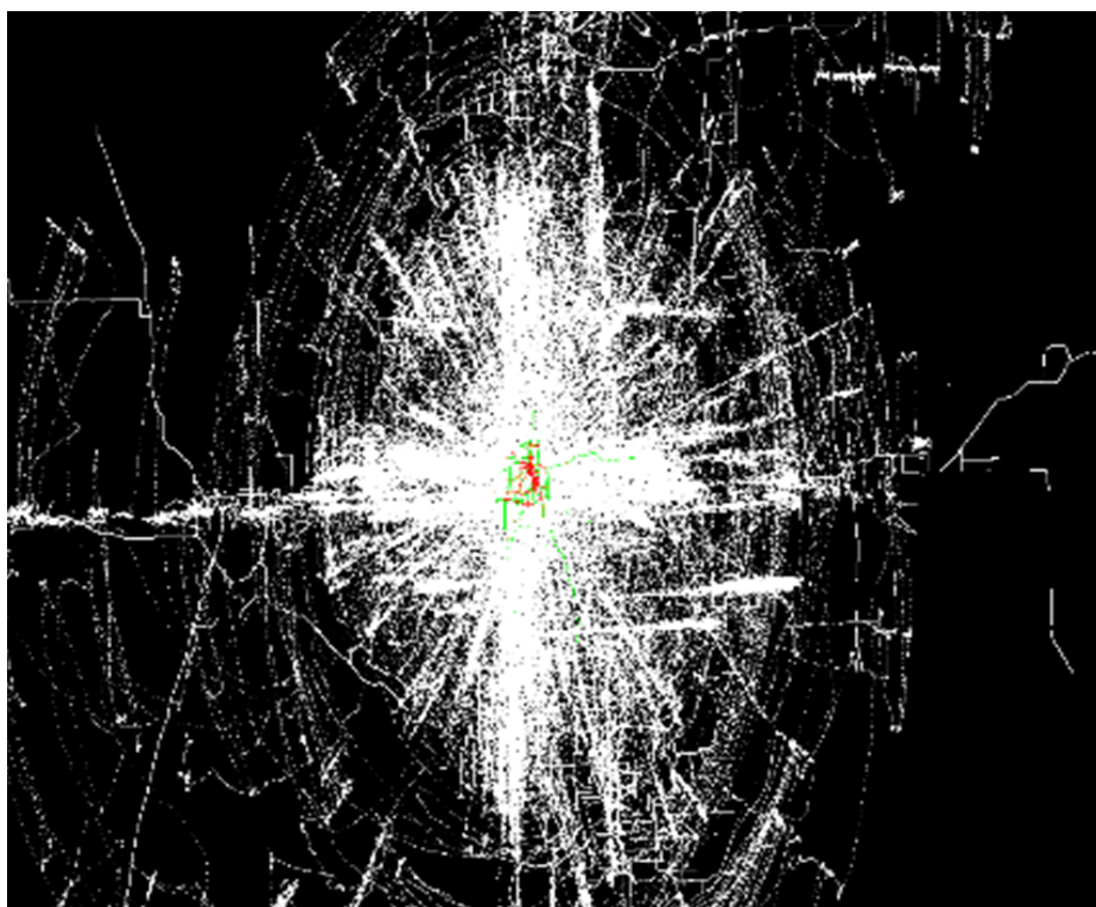


Рис. 5. Визуализация плотности движения транспортных средств с устройствами для передачи данных

Нечеткие срезы (fuzzy slices) – это хороший пример обогащения одной технологии (базы данных) другой (нечеткая логика). Под нечёткими срезами понимаются фильтры по измерениям, в которых фигурируют нечеткие величины, например, «все объекты, движущиеся на север города». В данном примере само понятие «север города» является нечетким, а если учесть, что объекты могут двигаться на север города не по прямой траектории, то нечеткость появляется в определении движения.

Для построения правил передачи данных через сеть DTN необходимо сделать несколько нечетких срезов исходной базы данных: по времени, по направлению и по расстоянию.

Закрашенная точка в центре – это выходной объект. Еще две закрашенные точки ближе к окружности – это движущиеся объекты со степенью принадлежности больше 0.8.

Результаты первого среза можно схематично изобразить в виде окружности радиусом 500 метров и некоторой совокупности точек, чье время и направление удовлетворяет заданным условиям (рис. 6).

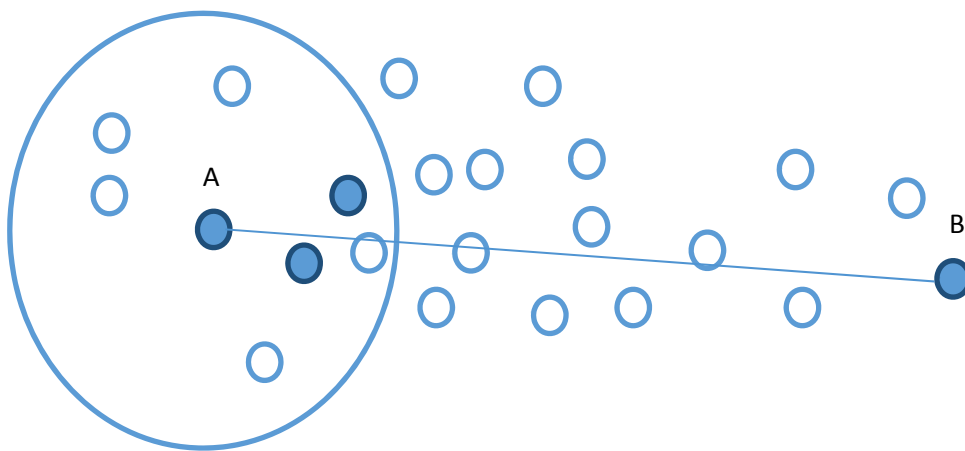


Рис. 6. Результаты первого среза

Повторяя эти срезы для найденных двух движущихся объектов, полагая каждый из них по очереди за входной объект будет построена целая совокупность окружностей (рис. 7).

Последними будут те окружности, которые содержат выходной объект. Алгоритм проведения срезов является конечным, так как начиная с определенной итерации, расстояние между входным объектом и выходным будет сокращаться вместе с числом движущихся объектов, обладающих степенью принадлежности больше 0.8.

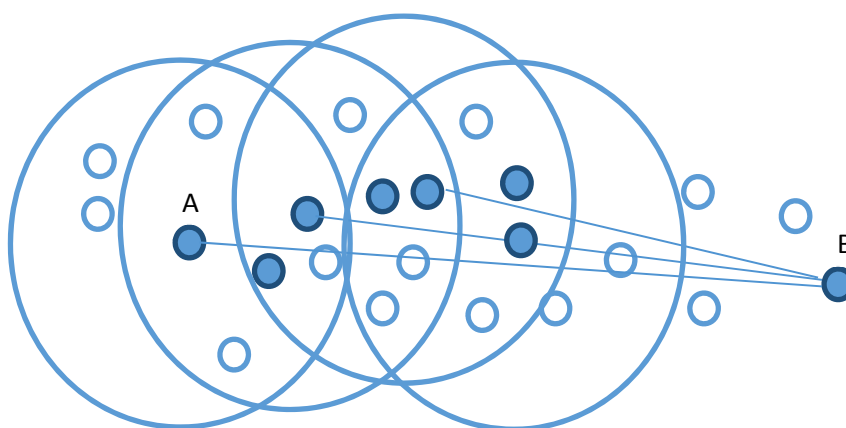


Рис. 7. Результаты следующих срезов

В результате работы всего алгоритма, основанного на нечетких срезах, будет получена совокупность связанных между собой точек, соответствующих движущимся объектам. Это множество точек по существу представляет собой связанный ориентированный граф с одной входной вершиной и одной выходной вершиной.

```
> use moto
switched to db moto
> use moto
switched to db moto
> use moto
switched to db moto
> db.preMesh_traffic_1.count()
387504024
> db.preMesh_traffic_2.count()
396409188
> db.preMesh_traffic_3.count()
360181073
> db.preMesh_traffic_4.count()
334144522
> db.preMesh_traffic_5.count()
342934063
> db.preMesh_traffic_6.count()
241741346
> db.preMesh_traffic_7.count()
204248857
>

> use moto
switched to db moto
> use moto
switched to db moto
> use moto
switched to db moto
> db.mesh_1.count()
1047
> db.mesh_2.count()
1015
> db.mesh_3.count()
1563
> db.mesh_4.count()
1188
> db.mesh_5.count()
1117
> db.mesh_6.count()
938
> db.mesh_7.count()
1957
>
```

Рис. 8. Вывод количества записей до и после применения нечетких срезов

Но, на практике в виду того, что данных оказалось очень много, то время выполнения таких срезов получилось значительным. На компьютере Intel(R) Core(TM) i3-4030U CPU 1.90GHz данные срезы по 2 267 163 158 записям с построением маршрута выполнялись бы

больше месяца. Поэтому число срезов было сокращено до двух: только по времени и по расстоянию. В результате число записей было значительно уменьшено до 7089 записей. Скриншоты по запросу объемов данных по дням недели приведены на рис. 8 и более визуально представлены на рис. 9.

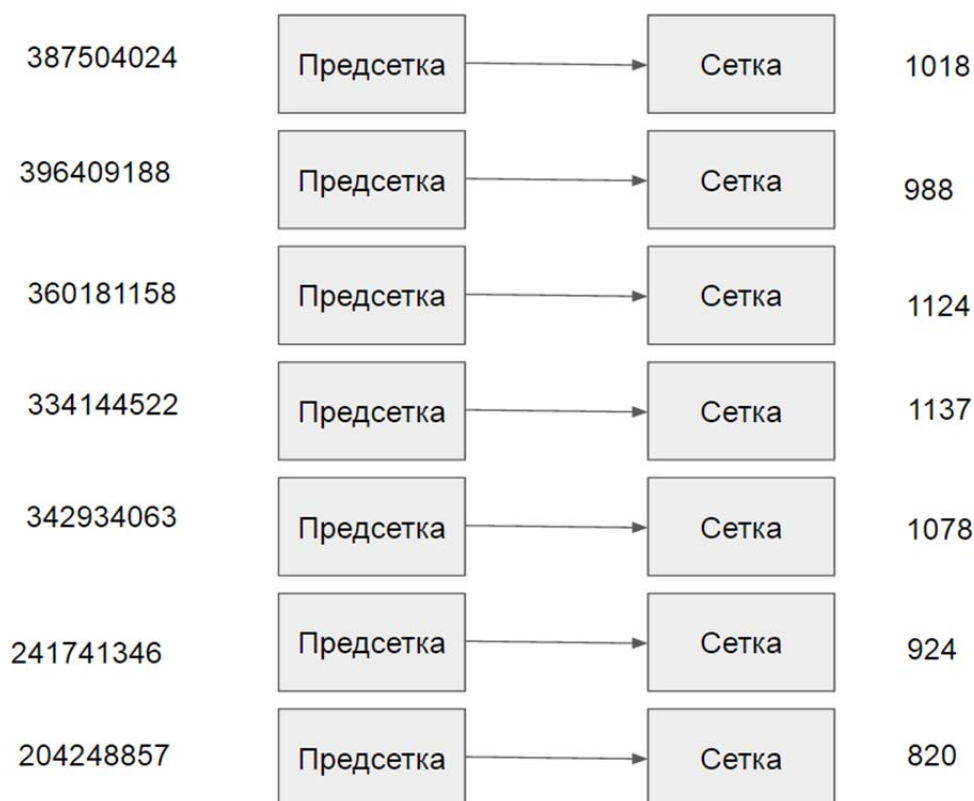


Рис. 9. Уменьшение числа записей после отсеивания ненадежных транспортных средств

На рис. 10 представлена координатная сетка с прореженными данными с помощью применения нечетких срезов.

Соответственно между ячейками координатной сетки был построен связный граф с учетом возможных направлений передачи информации (рис. 11).

К полученному графу был применен впоследствии алгоритма поиска в глубину и модифицированный алгоритм Йен для построения маршрутов. Первый из них получает JSON-строки из базы данных и строит по ним в зависимости от выбранных весов основной маршрут. С помощью модифицированного алгоритма Йена, в котором удаляются не каналы связи, а узлы, строятся альтернативные маршруты.

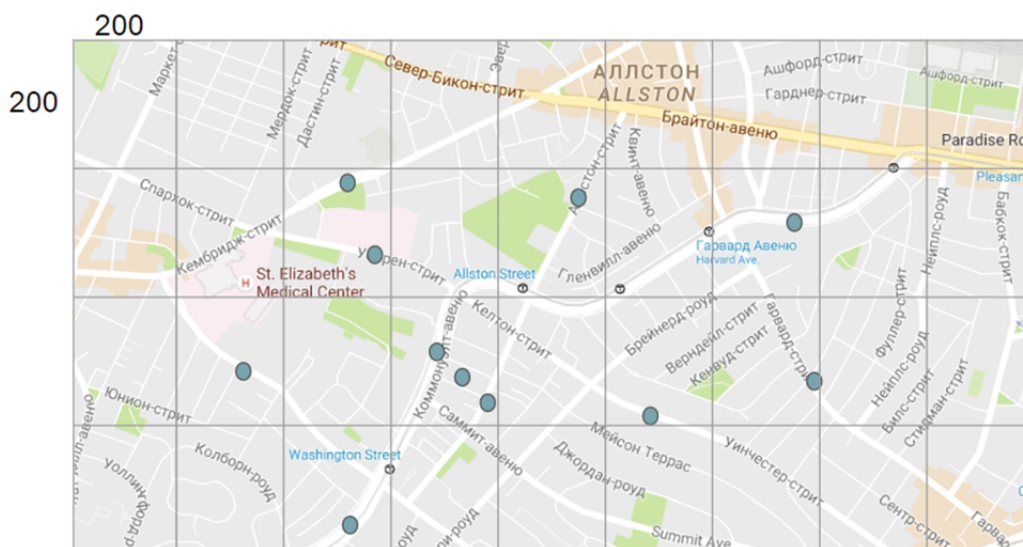


Рис. 10. Сетка после удаления из нее информации об устройствах, вероятность передачи данных через которые мала

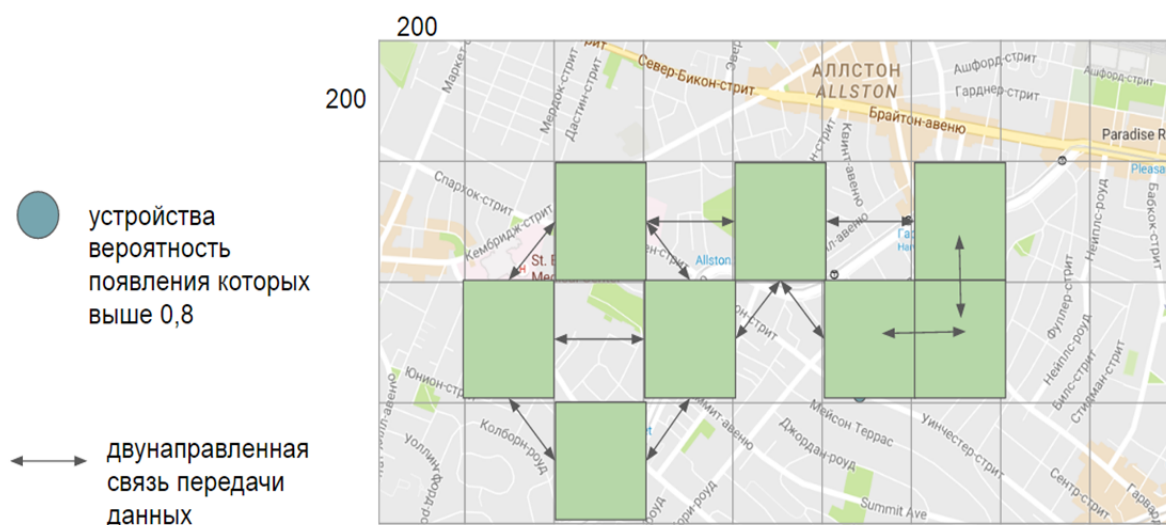


Рис. 11. Связный граф, построенный по укрупненной сетке

Предложенный подход к трансформации и отчистке данных позволяет сократить, в нашем случае, количество записей более чем в 256 000 раз.

Такое осязаемое уменьшение количества записей позволяет не только уменьшить затраты на хранение подготовленных данных, но и намного ускорить применение методов извлечения знаний. А применение библиотеки pthreads можно заставить РНР-приложения бережливо использовать процессорное время и получить выигрыш во времени выполнения

реализации решения задачи предварительной очистки и трансформации данных более чем в 2 раза на достаточно слабом процессоре.

В результате проведенных исследований было установлено, что существующие протоколы маршрутизации данных DTN–сетей могут быть разделены на три основные категории [21]: одна копия (или направленная пересылка), множество копий (или широковещательная рассылка) и гибрид (ограниченная широковещательная рассылка). Протоколы первого вида передают только одну копию сообщения по отобранному маршруту к месту назначения. Протоколы широковещательной рассылки передают множество копий сообщения к сенсорным узлам в пределах сети в ожидании, что, как минимум, одна копия достигнет места назначения. Протоколы широковещательной рассылки, например, протоколы, основанные на теории эпидемии или даже пандемии, [22], могут улучшить гарантии доставки. Однако, очевидно, при потерях связности сети и/или случайных разрывах соединений требования к объемам буферной памяти быстро растут при увеличении размера сети. Возникает следовательно необходимость поиска новых методов построения маршрутов в динамических сетях.

В данной статье представлены результаты исследований по построению маршрутов с применением нечеткой логики и теории графов. Показано, как с помощью нечеткой логики можно значительно уменьшить объем данных из предистории передвижения транспортных средств. Продемонстрировано как можно строить маршруты в два этапа. Сначала строить маршрут между множествами узлов. А впоследствии строить уточненный маршрут непосредственно уже между устройствами.

Рассмотренный в статье подход к построению маршрутов не подходит для транспортных средств, движущихся постоянно по новым маршрутам. Но, с помощью этого подхода можно исключить те транспортные средства, которые не имеют относительно постоянного маршрута.

9.6. Параллельная реализация процесса построения самоорганизующейся системы обмена информацией

Выше было показано, что работа самоорганизующейся системы обмена информацией, как и любой современной информационной системой, связана с обработкой большого объема данных в очень сжа-

тые сроки. Естественным решением задачи ускорения процессов обработки данных сегодня является применение технологий распределенных или параллельных вычислений.

В задаче генерации маршрутов в динамических сетях DTN для ускорения обработки запросов к базе данных об передвигающихся устройствах применялись методы разделения запросов на потоки с помощью библиотеки pthreads системы РНР.

В целом, разнообразие языков программирования дает возможность не только подобрать решение для конкретной задачи и получить максимальную отдачу от оборудования, но и выгоду в стоимости разработки и поддержки создаваемых систем.

РНР – один из наиболее популярных языков для создания динамических веб сайтов. Согласно рейтингу языков программирования, опубликованного GitHub [23] и основанного на количестве pull-запросов пользователей сервиса, РНР занимает пятую строчку, уступая JavaScript, Java, Python и Ruby. Исследования, проведенные аналитическим агентством «Тэглайн» [24], основанные на анкетировании 445 компаний, работающих в сфере цифровых технологий, также ставят РНР на первое место, так как 95% респондентов дали положительный ответ на вопрос о знании этого языка программирования.

Такая популярность обусловлена так называемым «низким порогом вхождения» и является огромной проблемой, так как многие начинающие разработчики даже не имеющие специального технического образования выбирают именно РНР как свой первый язык программирования. Именно это дало РНР нелестную славу инструмента, пригодного только для создания простейших сайтов и не способного на большее.

Но зачастую, востребованность в высоконагруженных решениях возникает не с самого начала. Переход на радикально иные технологии в больших проектах требует огромных затрат времени и материальных средств, что не всегда возможно, а у малых проектных групп может не быть финансовой возможности нанять разработчика на Java или С для реализации достаточно сложного функционала. В таких ситуациях приходится работать с тем, что имеется.

Ниже показано, как с помощью РНР можно реализовать высоконагруженное решение, реализующее процесс очистки и трансформации данных для подготовки к дальнейшему анализу при помощи методов Data Mining. Тестировании решения проведено на тестовых наборах в 25 000 и 1 000 000 записей

Наверное, не существует сегодня компаний, которые бы не собирали информацию в рамках своей деятельности. Одни ее собирают для отчетности, другие для интереса, но большинство для анализа и принятия решений, совершенствующих показатели компании. Пока информации немного работники компаний перерабатывают ее вручную, часто используя для этого табличные процессоры, запросы к базам данных и другие программные средства. Но, если поток поступающей информации значительный, то обрабатывать информацию становится с каждым разом все сложнее. Это занимает очень много времени. В этих случаях применяется специальная методика извлечения знаний из баз данных (Knowledge Discovery in Databases) [25]. Эта методика была предложена во второй половине двадцатого века Григорием Пятецким-Шапиро который задался вопросом автоматизации извлечения знаний из больших баз данных.

Концепция методики извлечения знаний не предлагает конкретных алгоритмов, но представляет порядок действий, которые необходимо выполнить для достижения поставленной цели и включает в себя:

- подготовку и сбор данных из источников;
- очистку данных от аномальных значений, пропусков и шумов и показателей, которые не несут полезной нагрузки в рамках поставленной задачи для снижения объема информации и размерности;
- трансформацию данных для приведения данных к виду, пригодному для извлечения знаний;
- собственно извлечение знаний путем применения различных алгоритмов;
- интерпретацию полученных знаний.

Рассмотрим задачу: извлечении полезной информации из базы, в которой находятся данные, полученные с устройств, передающих свои GPS-координаты.

Для решения этой задачи для ускоренного применения конечных методов Data Mining была произведена предварительная очистка и трансформация данных, реализованная на PHP и более глубокая очистка и трансформация данных путем применения нечетких срезов баз данных, реализованная на языке Java. В качестве хранилища информации была использована noSQL база данных MongoDB версии 3.2.

Более подробно представленные далее процессы очистки и трансформации данных, а также результаты применения нечетких срезов в базах данных в качестве метода глубокой очистки описаны в [26].

Задача, рассматриваемая в этой статье, является продолжением задачи об извлечении полезной информации из базы, и заключается в следующем: провести исследования производительности РНР7.0 во время предварительной очистки и трансформации данных, определить каким способом можно значительно сократить время предварительной очистки данных.

9.6.1. Многопоточность в РНР. Библиотека `pthread`s

Хотя сам по себе РНР и не поддерживает многопоточность, есть несколько методов её эмуляции. Самый простой – это запуск нескольких копий скрипта с параметрами. Такой подход неприемлем в высоконагруженных системах и задачах, так как не позволяет контролировать выполнение каждого экземпляра, выделение памяти и ресурсов.

Следующий подход – PCNTL. Это расширение позволяет полноценно работать с потоками и является довольно простым для тех, кто освоил многопоточные вычисления в языке С. При использовании расширения PCNTL разработчику доступна функция `pcntl_fork`, которая является аналогом `fork` в С и позволяет порождать дочерние процессы. Но опять же, этот подход является имитацией многопоточности и не так прост для неподготовленного программиста.

Joe Watkins, автор расширения `pthread`s, предлагает объектно-ориентированное API для реализации по-настоящему многопоточных вычислений [27]. `Pthreads` позволяет приложению выполнять многопоточные вычисления и синхронизировать потоки между собой. Это расширение включает в себя несколько классов:

- `Threaded` – главный класс в `pthread`s. Именно благодаря в РНР появляется возможность параллельного запуска потоков. В нем заключен метод `synchronized()` для синхронизации и многие другие методы.

- `Thread` – создает отдельный поток путем вызова метода `run()`. Метод `run()` начинает выполняться в отдельном потоке в момент вызова метода `start()`. Это можно инициировать только из контекста, который создает поток. Объединить потоки можно тоже только в этом же контексте.

- `Worker` – это по сути, исполнитель потока.

- `Pool` – пул-контейнер из нескольких `Worker`. По мере освобождения очередного `Worker` запускается следующий экземпляр `Threaded`.

Таким образом, pthreads – является чуть-ли не единственным способом «безболезненно» организовать многопоточные вычисления в PHP.

Следует отметить, что для запуска pthreads необходимо скомпилировать PHP с опцией `--enable-maintainer-zts` (Zend Thread Safety), но для опытного администратора это не составит проблем, а для пользователей Debian-based Linux систем в репозитории `ppa:ondrej/php-zts` есть два пакета `php7.0-zts` и `php7.0-zts-dev`, которые добавляют в систему `php7.0-zts`, который можно запустить из терминала.

9.6.2. Решение задачи предварительной очистки и трансформации данных без использования библиотеки pthreads

Накопленные знания о передающих GPS-устройствах представляют собой журналируемую информацию об устройстве, в которой вместе с координатами, содержатся показатели ускорения, вращения, скорости, данные о передающем и принимающем устройствах, а также и другая служебная информация. Такой набор данных избыточен для поставленной задачи и не имеет некоторых производных от существующих показателей. Чтобы в дальнейшем возможно было применить методы извлечения знаний имеющуюся информацию необходимо подготовить.

Предварительная очистка и трансформация в рамках нашей задачи заключается в следующем:

- На этапе очистки необходимо произвести исключение из обработки «мусорной» информации и данных, которые были получены путем некорректного получения сигнала от GPS-спутников. Эти данные географически расположены довольно далеко от остального набора и могут быть исключены из обработки путем вычисления расстояния до ближайших записей из отсортированного по индексированным полям массива.

- На этапе трансформации необходимо сделать проход по по двумерному массиву географических координат и выполнить привязку данных, относящихся к определенному сектору географической координатной плоскости к точке в новом координатном пространстве. Получение из «сырых» данных новых, полезных для дальнейшей обработки возможно путем выполнения простых арифметических операций.

Данный набор операций был реализован с использованием PHP 7.0 с драйвером `mongodb-1.2` без использования библиотеки `RHPLIB` для MongoDB на тестовом стенде с характеристиками, приведенными в таблице 1.

Таблица 1

Характеристики тестового стенда

Процессор	
Model name:	Intel(R) Core(TM) i3-4030U CPU @ 1.90GHz
Socket:	1
Core(s) per socket:	2
Thread(s) per core:	2
Память	
RAM:	12 Gb
Storage:	120 Gb (SSD)
Операционная система	
Description:	Linux Mint 18 Sarah
Linux Kernel:	4.4.0-21-generic

Одной из основных проблем, связанных с обработкой больших данных является получение максимальной отдачи от оборудования, на котором выполняется обработка информации. Не смотря на то, что большинство баз данных легко масштабируются для работы в несколько потоков, исполнение скриптов на РНР, как и на других языках, без дополнительной подготовки будет исполняться в один поток. При таком подходе невозможно получить от конечного оборудования максимальной отдачи и график загрузки логических ядер будет выглядеть приблизительно так:

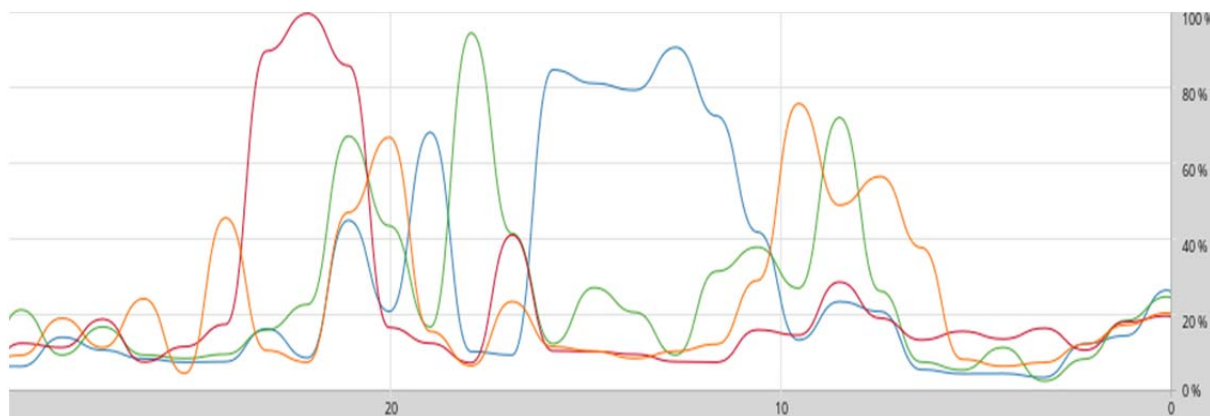


Рис. 12. График загрузки логических ядер на наборе из 35 000 записей

Как видно из графика, в процессе обработки данных задействованы все потоки. Это обусловлено тем, что в MongoDB начиная с версии 3.0 используется новый механизм хранения WiredTiger, который

работает по всем доступным потокам, но ни один из потоков не загружен на максимальную производительность, что говорит о неэффективном использовании имеющихся ресурсов.

Одним из путей решения данной задачи является запуск процесса обработки параллельно в нескольких потоках. Именно для таких случаев в РНР есть несколько подходов для реализации многопоточных вычислений.

9.6.3. Параллельная реализация решения задачи предварительной очистки и трансформации данных с использованием библиотеки `pthread`s

В процессе реализации многопоточных вычислений возникает вопрос, как и в каком месте кода реализовать ветвление в потоки и как обеспечить синхронизацию этих вычислений. Так как библиотека `pthread`s предоставляет метод `synchronized()` в классе `Threaded`, остается решить, в каком моменте возможно запустить отдельный поток.

В рамках решения задачи предварительной очистки и трансформации данных такими местами были:

- процесс очистки от аномальных значений по каждой из координат (по долготе и широте). В данном случае есть возможность запустить два потока (по одному на каждую из координат)
- обход координатного пространства по долготе и по широте, когда реализуется вложенный цикл `for`. Внутренний цикл можно запустить в отдельном потоке.

После определения таких участков можно приступить к реализации параллельных вычислений. Например, если ранее в последовательной реализации решения задачи предварительной очистки и трансформации данных было два вложенных цикла:

```
for ($long=$this->minLong; $long<=$this->maxLong; $long+=$this->dLong) {
    for ($lat = $this->minLat; $lat <= $this->maxLat; $lat += $this->dLat) {
        // Далее вызывается драйвер MongoDB для поиска записей,
        // входящих в рассматриваемый сектор.
        // К полученным записям применяется процесс трансформации.
    }
}
```

То после реализации этого решения вместе с библиотекой pthreads получим:

```
public function run()
{
do {
    $val = null;
    // Получаем объект, наследуемый от Threaded
    $prov = $this->worker->getProvider();
    // Синхронизация данных
    $prov->synchronized(function($prov) use (&$val) {
        //Метод getNext() реализует замену внешнего цикла
        $val = $prov->getNext();
    }, $prov);
    if ($val['next'] === null) {
        continue;
    }
    // Метод transformation() реализует замену внутреннего цикла и
    // выполняется параллельно
    $this->transformation($val);
}
while ($val !== null);
}
```

В результате исследований было установлено, что использование библиотеки pthreads рационально для реализации высоконагруженных систем и вычислений.

На Рисунке 12 была показана диаграмма загрузки потоков при использовании PHP без ZTS и pthreads. Далее, будут показаны диаграммы загрузки логических ядер на процессоре в зависимости от количества запущенных потоков. Можно заметить, что диаграмма, полученная при выполнении последовательного решения задачи (Рисунок 1) сравнительно схожа с диаграммой, полученной при использовании pthreads с 1 открытым потоком (рис. 13).

При открытии одного дополнительного потока нагрузка распределяется гораздо лучше, чем в однопоточной реализации, но на диаграмме (рис. 14) заметно, что постоянная пиковая загрузка не реализована.

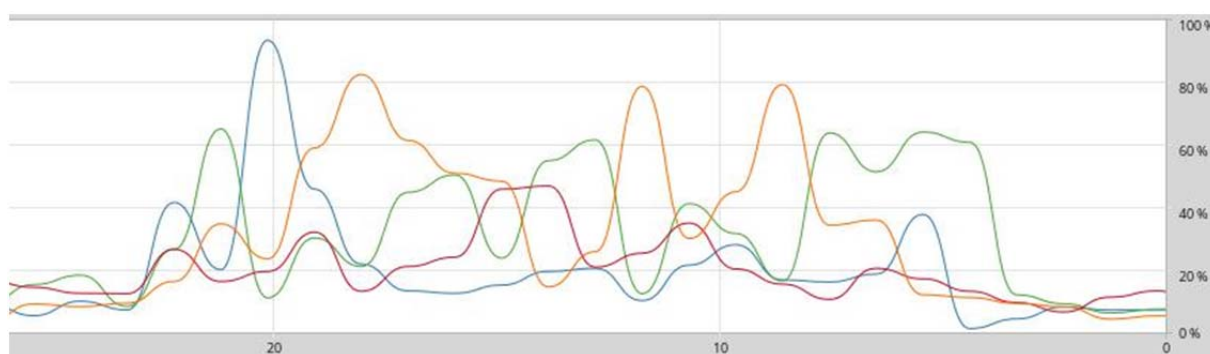


Рис. 13. График загрузки логических ядер при использовании pthreads с 1 открытым потоком на наборе из 35 000 записей

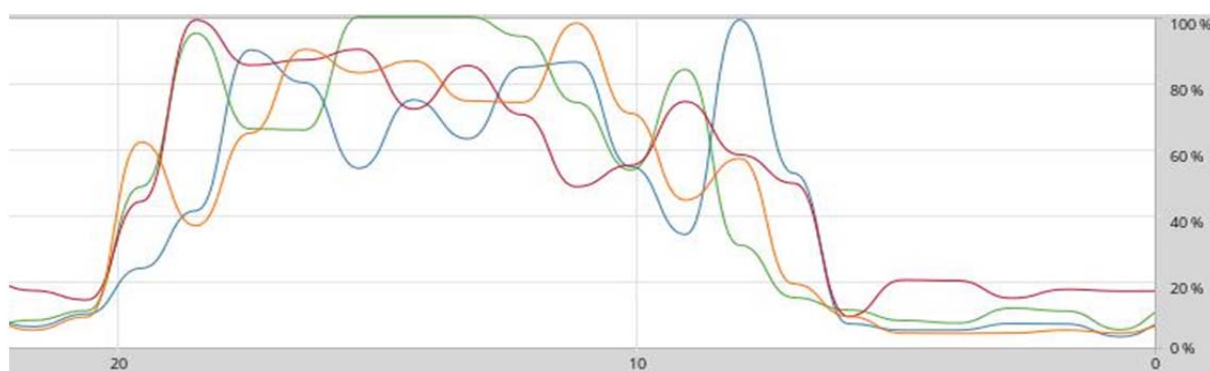


Рис. 14. График загрузки логических ядер при использовании pthreads с 2 открытыми потоками на наборе из 35 000 записей

Загрузка с тремя открытыми потоками (рис. 15) выводит среднюю загрузку процессора на уровень 80%. Такой уровень оставляет простор для использования вычислительных мощностей другими пользователями с менее требовательными задачами, но не реализует в полной мере всего потенциала для вычислений.

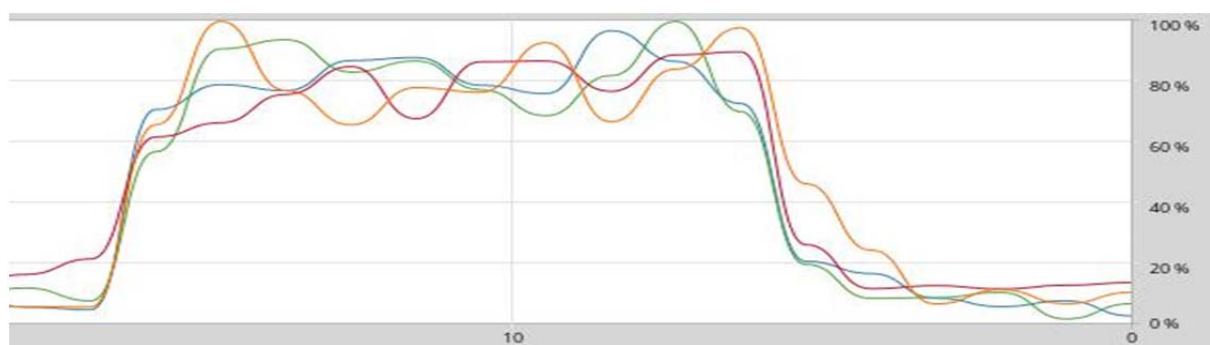


Рис. 15. График загрузки логических ядер при использовании pthreads с 3 открытыми потоками на наборе из 35 000 записей

При открытии дополнительных потоков становится заметно, что загрузка логических ядер растет, и достигает своего пика при использовании всех доступных потоков (рис. 16). Такие показатели говорят о том, что при помощи расширения pthreads можно заставить РНР работать достаточно производительного не только для web-приложений, но и для решения задач, связанных с извлечением знаний из больших данных.

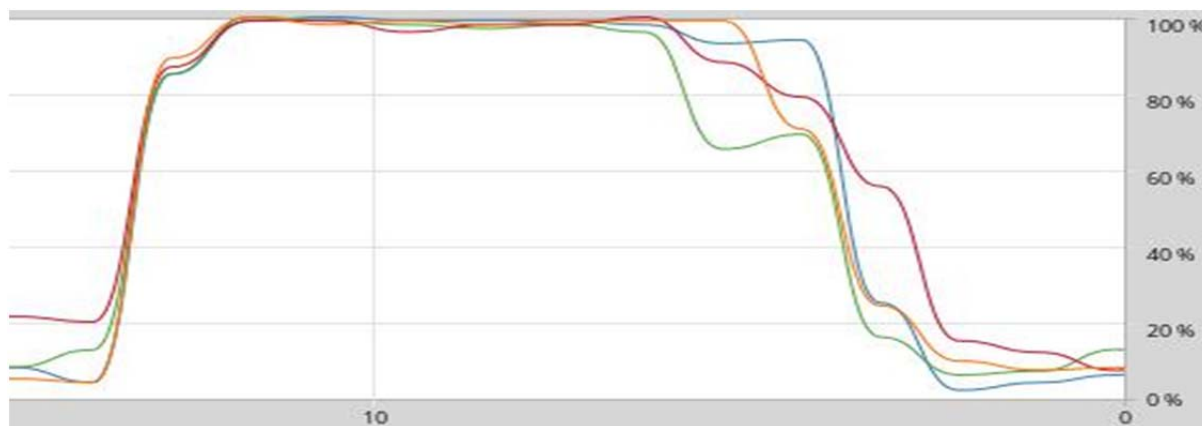


Рис. 16. График загрузки логических ядер при использовании pthreads с 4 открытыми потоками на наборе из 35 000 записей

Для сравнения замерим время выполнения последовательного и параллельного решения задачи предварительной очистки и трансформации данных на тестовых наборах. Для проведения более точных вычислений на каждом наборе выполним по 100 замеров, а средние показатели среди них занесем в таблицу, представленную ниже.

Таблица 2

Результаты замеров времени работы решения задачи

Алгоритм	Время выполнения, сек.	
	25 000 записей	1 000 000 записей
Последовательный (NTS)	17,09	233,85
Последовательный (ZTS)	16,99	228,97
Параллельный в 1 поток	17,00	224,94
Параллельный в 2 потока	10,01	127,32
Параллельный в 3 потока	9,12	98,67
параллельный в 4 потока	8,27	79,67

Полученные результаты показывают что время выполнения последовательного решения на стандартной сборке РНР7.0 без исполь-

зования Zend Thread Safety (NTS) практически не отличается от последовательного решения задачи, запущенного на РНР с использованием Zend Thread Safety (ZTS) как и время решения, выполненного с использованием pthreads и открытием всего одного потока. Однако, при открытии более одного потока получаем колоссальный прирост в производительности, благодаря которому время выполнения задачи уменьшается со скоростью кубической регрессии. Рисунок 17 демонстрирует зависимость времени выполнения решения от количества открытых потоков.

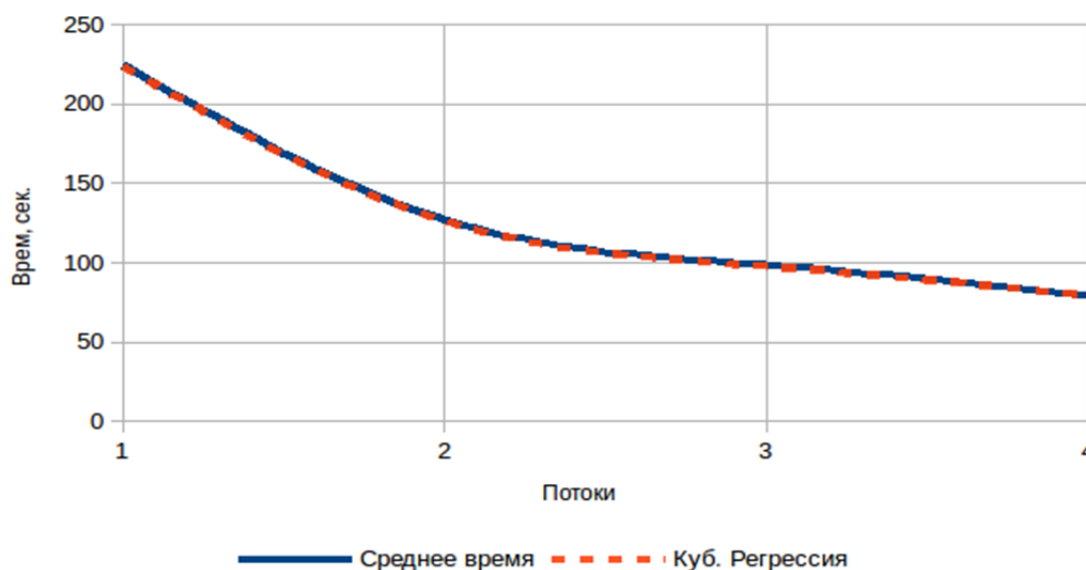


Рис. 17. График зависимости времени выполнения процесса решения задачи от количества открытых потоков на наборе из 1 000 000 записей

Полученные результаты показывают, что применяя библиотеку pthreads можно заставить РНР-приложения бережливо использовать процессорное время и получить выигрыш во времени выполнения реализации решения задачи предварительной очистки и трансформации данных более чем в 2 раза на достаточно слабом процессоре.

Список литературы

1. *Винер Н.* Кибернетика, или управление и связь в животном и машине. М.: Сов. Радио, 1968. 326 с.
2. *Ивахненко А.Г. и др.* Принятие решений на основе самоорганизации. М.: Сов. Наука, 1978.

3. *Кузнецова В.Л., Раков М.А.* Самоорганизация в технических системах. Киев: Наук. Думка, 1987. С. 200.

4. *Холод И.И.* Самонастраивающиеся агенты для интеллектуального анализа распределенных данных. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 1. С. 95–96.

5. Самонастраивающиеся агенты для интеллектуального анализа распределенных данных / Холод И.И. и др. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 1. С. 95–96.

6. *Чирков А.Б.* Карты Кохонена и принцип максимизации взаимной информации. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 2. С. 155–157.

7. *Гнеденко Б.В., Коваленко Н.Н.* Введение в теорию массового обслуживания. М.: Наука, 1966. 240 с.

8. *Писаревский А.С.* Системы с перестраиваемой архитектурой. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям // Известия СПбГЭТУ «ЛЭТИ» 2006. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. С. 46–49.

9. *Абакумов Р.С., Куприянов М.С., Писаревский А.С.* Методы реконфигурации FPGA-структур. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2005. № 3. С. 72–76.

10. *Чирков А.Б.* Применение самоорганизующейся карты Кохонена для решения задачи кластеризации // Известия СПбГЭТУ «ЛЭТИ» 2006. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. – С. 90–95.

11. *Куприянов М.С., Кочетков А.В.* Методы определения консистентного состояния в системах с самовосстановлением // Известия СПбГЭТУ «ЛЭТИ». СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. С. 78–85.

12. *Кочетков А.В.* Разработка алгоритмов самовосстановления распределенных систем. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 1. С. 43–45.

13. *Елизаров С.И., Куприянов М.С., Холод И.И.* Самонастраивающиеся агенты для интеллектуального анализа распределенных данных // Известия СПбГЭТУ «ЛЭТИ». № 2. 2006. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. С. 3–7.

14. *Куприянов М.С.* Методы самовосстановления информационных систем. Сб. докл.: Международная конференция по Мягким вы-

числениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т 1. С. 189–194.

15. Кадамов А.Д., Куприянов М.С., Маркин А.В., Лукашевич С.В., Тревгода М.А., Тринкунас А.В., Пашковский С.В. Генерация тестовых сценариев для USE-CASE модели. Сб. докл.: Международная конференция по Мягким вычислениям и измерениям. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006. Т. 1. С. 207–209.

16. Орловский С.А. Проблемы принятия решений при нечёткой исходной информации. М.: Наука, 1981.

17. Bolodurina I., Parfenov D., Shukhman A. Approach to the effective controlling cloud computing resources in data centers for providing multimedia services // Control and Communications (SIBCON), 2015 International Siberian Conference on. – IEEE, 2015. P. 1–6.

18. Charuenporn P., Intakosum S. Qos-Security Metrics Based on ITIL and COBIT Standard for Measurement Web Services // J. UCS. 2012. Vol. 18. № 6. P. 775–797.

19. DAndria F. et al. Cloud4soa: multi-cloud application management across PaaS offerings // Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2012 14th International Symposium on. – IEEE, 2012. P. 407–414.

20. Zhu J., Zheng Z., Lyu M.R. DR2: Dynamic Request Routing for Tolerating Latency Variability in Online Cloud Applications // IEEE CLOUD. 2013. P. 589–596.

21. Tanenbaum, A.S. Computer Networks, 5th Ed. [Text] / Andrew S. Tanenbaum, David J. Wetherall. Prentice Hall, Cloth, 2011. 960 p.

22. Mundur, P. Delay Tolerant Network Routing: Beyond Epidemic Routing [Text] / P. Mundur., M. Seligman // Proceedings of third international symposium on Wireless Pervasive Computing, 2008. – ISWPC 2008, 7–9 May 2008. P. 550–553.

23. The state of the Octoverse 2016 [Электронный ресурс] // GitHub Octoverse 2016. URL: <https://octoverse.github.com/> (дата обращения: 01.03.2017).

24. Рейтинг языков программирования 2016 [Электронный ресурс] // Тэглайн – свежие рейтинги и исследования Рунета. 2016. 11 апреля. URL: <http://tagline.ru/programming-languages-rating/> (дата обращения: 01.03.2017).

25. Piatetsky-Shapiro G., Frawley W. Knowledge Discovery In Databases // AAAI Press. – December 1991. ISBN 9780262660709. 539 p.

26. *Yulia Shichkina, Alexander Degtyarev, Alexander Koblov*. Technology of cleaning and transforming data using the Knowledge Discovery in Databases (KDD) technology for fast application of Data Mining methods // CEUR Workshop proceedings. Selected Papers of the 7th International Conference Distributed Computing and Grid-technologies in Science and Education. 2017. Vol. 1787. Pp. 428–434. urn:nbn:de:0074-1787-5.

27. PHP: pthreads – Manual [Электронный ресурс] // PHP: PHP Manual. URL: <http://php.net/manual/en/book.pthreads.php> (дата обращения: 01.03.2017).

ИНФОРМАЦИЯ ОБ АВТОРАХ

Ю.В. Барышев – д.ф.-м.н., вед. науч. сотр., Санкт-Петербургский Государственный университет, кафедра астрономии.

В.В. Витковский – к.ф.-м.н., профессор, Санкт-Петербургский национальный исследовательский университет информационных технологий механики и оптики кафедра астроприборостроения.

В.Л. Горохов – д.т.н., профессор, Санкт-Петербургский национальный исследовательский университет информационных технологий механики и оптики, кафедра астроприборостроения, **САО**.

М.С. Куприянов –

В.Л. Лазарев – Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики» (ITMO University), holod25@yandex.ru

Л.Г. Логинская –

С.В. Прокопчина –

В.Б. Тарасов –

Ю.А. Шичкина –

CONTENTS

Preface	10
Chapter 1. Soft approaches to the measurement and management of complex systems. The regularizing Bayesian approach	12
1.1. Концептуальные основы решения задач в условиях неопределенности	12
1.1.1. Basic principles of solving problems of measurement and control of complex systems in conditions of uncertainty	12
1.1.2. Обобщенная методика и схема решения задач в условиях неопределенности	19
1.1.3. Методологические основы регуляризирующего байесовского подхода для решения задач в условиях неопределенности	26
1.1.4. Модели и шкалы с динамическими ограничениями. Концепция, типы, структура	32
1.1.5. Основные принципы интеграции информации на основе регуляризирующего байесовского подхода (РБП) и его технологий	41
1.1.6. Метрологическое обоснование информационных технологий решения задач в условиях неопределенности	45
1.2. Организация измерительных процессов в условиях неопределенности. Регуляризирующий байесовский подход	47
1.2.1. Исторические аспекты развития теории измерений и интеллектуализации измерительных процессов	47
1.2.2. Основная концепция байесовской интеллектуализации измерений	55
1.2.3. Новые типы измерений в условиях неопределенности: нечеткие, мягкие, когнитивные измерения	56

1.2.4. Когнитивные измерения на основе байесовских интеллектуальных технологий	63
1.3. Интеллектуальные информационные технологии на основе регуляризующего байесовского подхода	73
1.3.1. Интегрирующие технологии на основе байесовских интеллектуальных измерений	73
1.3.2. Принципы создания развивающихся информационных технологий на основе регуляризующего байесовского подхода	76
1.3.3. Интеллектуальные технологии решения многокритериальных задач в условиях риска и неопределенности на основе регуляризующего байесовского подхода	83
1.3.4. Методология и технологии мягких критериев и мягкого нормирования	92
1.3.5. Мягкие функции и их применение в управлении сложными системами	95
1.3.6. Методология и технологии байесовской математической статистики на основе РБП. Система математической обработки статистической информации «байесовская математическая статистика»	99
1.4. Мягкие технологии аудита и управления сложными системами в условиях неопределенности	106
1.5. Решение прикладных задач на основе регуляризующего байесовского подхода	122
1.5.1. Опыт применения РБП и БИТ для решения измерительных и управленческих задач	122
1.5.2. Реализация принципов экономики знаний на основе БИИ	125
1.5.3. Решение социально-экономических задач на основе байесовских интеллектуальных технологий	131
1.5.4. Моделирование банковского сектора и показателей экономических систем ЕАЭС	133
1.5.5. Методологические аспекты создания интеллектуальных энергетических сетей (smart grid) .	137
Literature	142
Chapter 2. From fuzzy sets to soft estimates and synergetic artificial intelligence	145

Introduction	145
2.1. NO-factors, mild estimates and engineering of opinions	147
2.2. Linguistic approach to the formalization of NO-factors:	
principles, stages, models	151
2.2.1. Simulation of fuzziness	152
2.2.2. Modeling the inconsistency	154
2.2.3. Simulation of inaccuracy	155
2.3. From families of NO factors to hybrid models	
and synergistic AI	155
Literature	157
Chapter 3. Methods of constructing granular logical values	
and structures	160
Introduction	160
3.1. Granular truth values and non-standard semantics	161
3.1.1. Classical interpretations of truth values	161
3.1.2. Non-standard logical semantics and generalized truth	
values	164
3.2. From logical worlds to granular logical matrices	166
3.2.1. Logical worlds and spaces	166
3.2.2. Logical matrices	168
Literature	171
Chapter 4. Modeling of opinions and assessments of intelligent	
agents: from four-valued systems of modalities to nonclassical	
measures and fuzzy allocations	174
Introduction	174
4.1. Opinions and modalities	176
4.1.1. Opinions	176
4.1.2. Descriptions and prescriptions in the activity of the	
agent	178
4.1.3. Values, assessments and norms	179
4.1.4. A unified approach to the description of various	
modality systems based on four-valued semantics	181
4.1.5. Lattice modalities	183
4.2. Nonclassical measures and description of modalities	184
4.2.1. Probabilistic measure	184
4.2.2. Family of measures Sugeno	185
4.2.3. Fundamentals of the theory of evidence	186
4.2.4. Measures of possibility and necessity	187

4.2.5. Qualitative assessments of the possibility and fuzziness	192
4.3. Measures on fuzzy sets	192
4.3.1. Measures of energy (indicators of strength of belonging)	193
4.3.2. Measures of entropy	194
4.3.3. Specificity measures	194
Conclusion	195
Literature	196
Chapter 5. The theory of entropy potentials. State, development prospects and practical applications	199
Introduction	199
5.1. Starting positions	200
5.2. Overview of approaches to describing uncertainty states	202
5.3. Fundamentals of the theory of entropy potentials	204
5.4. Study of systems based on portraits in the space of characteristics of entropy potentials	213
5.5. Information portraits	217
5.6. Application of TEP methods to solving practical problems	228
Conclusion	234
Literature	235
Chapter 6. Methodology of cognitive visualization of multidimensional data	240
Introduction	240
6.1. Cognitive technologies as the basis for a new generation of intelligent solution support systems	241
6.1.1. The current state of research on this issue, the main directions of research in the world science	241
6.1.2. Direction of research	241
6.1.3. List of main publications	244
6.2. Expanding and updating the field of rational data analysis based on cognitive interfaces	246
6.2.1. Mechanisms for visualization of multidimensional data	248
6.2.2. An example of cognitive visualization of multidimensional terabyte data of COSMOS astrophysical survey	253
6.3. An example of the detection of abnormal situations in the monitoring of complex systems based on cognitive technologies	254
6.3.1. Description of the procedure for detecting anomalies	255

6.3.2. Identification of dynamics	258
6.3.3. Identifying dependencies on data pairs	259
6.3.4. Comparison with other methods	261
6.4. Technology and methodology for generating pseudo-6-D images	262
6.5. Identification of extreme states of complex natural and technical systems based on cognitive monitoring of multidimensional time series	265
6.5.1. Algorithms of cognitive visualization of multidimensional time series	266
6.5.2. Practical application of cognitive technologies for detecting abnormal conditions of monitoring objects ...	267
6.6. Nonparametric statistics in multidimensional time series for cognitive detection of anomalies	269
6.6.1. Stages of cognitive analysis	269
6.6.2. Robust (robust) algorithms for detecting anomalies	270
6.6.3. Practical use	271
6.7. The practice of forming engineering visual metaphors on the basis of cognitive images of multidimensional data of technospheric monitoring	274
6.8. Cognitive technologies for visualization of multidimensional data for intellectual decision support	276
6.8.1. Principles of work	277
6.8.2. New algorithms of cognitive visualization	280
6.9. Semiotics of Visual Engineering Metaphors of Cognitive Images of multidimensional Data he paradigm of computer ergonomics	281
Conclusion	284
Literature	285
Chapter 7. Fuzzy methods of self-organization of information systems	291
7.1. Fuzzy numbers and logical values	292
7.2. Semantic scales for measuring the meanings of qualitative concepts	296
7.3. Procedure for generating a basis of fuzzy terms associated with a linguistic pair	296
7.4. Semantic scales for measuring the meaning of qualitative concepts	297
7.5. Decomposition of fuzzy value associated with the linguistic	299

pair according to the basic terms of this pair	
7.6. Approximation of Boolean Values	300
7.7. Measurement of qualitative quantities	301
7.8. Graduation of the linguistic scale	302
7.9. Arithmetic operations on fuzzy numbers	304
7.10. Logical variables and operations	307
7.11. Approximation of Boolean Values	308
7.12. Logical scale as a means of connecting different scales to each other	309
7.13. Approach to the definition of operations on fuzzy values given by functions with several maxima	310
7.14. Modification of values calculated by a fuzzy algorithm	311
7.15. Analysis of the efficiency of redundant systems based on qualitative information represented by fuzzy numbers	315
7.16. Structure of the scale for measuring the health of the unit	317
7.17. Analysis of the health of the blocks	317
7.18. Technique of interrogation of experts for clarifying the values of S_{ij} - the degree of involvement of the block B_j to the failure of M_i	320
7.19. The construction of s_{ij}	320
7.20. The algorithm for constructing s_{ij}	322
7.21. Failure interval forecast	325
7.22. Algorithm of forecast change	326
7.23. Estimation of the correctness of the forecast	327
7.24. Disabling forecasting processes	328
7.25. The current average time is	328
7.26. Determining the health of the block	329
7.27. Changing the value of the qualitative value of the current average	330
Literature	331
Chapter 8. Methods for increasing the efficiency of genetic algorithms	333
8.1. Theory of GA, features and basic dependencies	333
8.1.1. The process of optimization with the help of GA and the concept of circuits	333
8.1.2. The circuit theorem	339
8.1.2.1. Selection	340
8.1.2.2. Crossover	341
8.1.2.3. Mutation	341

8.1.3. Increasing the effectiveness of genetic algorithms	342
8.1.4. Development of a generalized scheme of GA	346
8.2. Research and development of the selection operator	349
8.3. Investigation of mutation and crossover operators	352
8.4. Development of fuzzy adaptive homogeneous crossover	353
8.4.1. The essence and methods of implementation	353
8.4.2. Development of the NBP for NAOK	355
8.4.3. Serial BHA + with NAOK	358
8.4.4. Results of experiments	363
Literature	368
Chapter 9. General concept and approach to the construction of distributed self-organizing information systems (ISIS)	370
9.1. Traditional approaches to building information systems	371
9.2. Classification of self-organizing systems	373
9.3. Mathematical model of self-organizing information system (SIS)	373
9.4. Structural and functional model of SIS	375
9.5. An example of building a self-organizing information exchange system	380
9.5.1. Setting the problem of generating rules in dynamic networks	382
9.5.2. Preliminary preparation of data	383
9.5.3. The method of extracting knowledge from databases. The concept of grid drawing	384
9.5.4. Visualization of the received data	387
9.5.5. Application of fuzzy logic to thin out data in a grid	388
9.6. Parallel implementation of the process of building a self- organizing information exchange system	393
9.6.1. Multithreading in PHP. The pthreads library	396
9.6.2. Solution of the task of preliminary cleaning and transformation of data without using the pthreads library	397
9.6.3. Parallel implementation of the solution of the pre- cleaning and data transformation task using the pthreads library	399
Literature	403
Information about authors	407

CHAPTER 1

Soft approaches to the measurement and management of complex systems. The regularizing Bayesian approach

S.V. Prokopchina

***Abstract.** The concept, methods, tools of Regularizing Bayesian Approach (RBA), Bayesian Intelligent technologies (BIT), soft measurement (SM), cognitive measurement (CM), polysystem's measurement (PM) as the new type of intelligent measurement, soft technologies, soft criteries (SC), soft functions (SF) have suggested. The methodological aspects to theoretical base creation are considered. The main examples of concept realization and technologies are given. The models with dynamic constraints, Scales with dynamic constraints are considered. The examples of BIT, SM, CM, PM, SC, SF technologies on the base of program platform «Infoanalyst» are developed. The practical examples are considered.*

The perspectives of information technologies development are considered. The concept of creation self-developing information technologies on the base of information technologies management, audit and monitoring is suggested. The realization of concept's principals on the base of Regularization Bayesian Approach is given. The examples of scale with of dynamics limitation usage for realization concept are shown. The general scheme of concept's ideas realization is suggested.

The software system of BMS (Bayes Mathematical Statistic) is described in this article. It is the system working with statistical data using different algorithms. This system allows constructing and executing different algorithms to calculate statistical characteristics and checking of statistical hypothesis. The system is intended for processing data on the base of Regularization Bayesian Approach that allows to apply it for decision of problems in conditions of significant uncertainty

The measurement of the potential and risk is considered.

The subsystem of RBE-Prognosis (Regularizing Bayesian Evaluation Prognosis) is described in this article. It is the prognosis system working with knowledge represented by RBE. This system is a part of the Infoanalyst – unique program to knowledge analysis under conditions of risk and uncertainty.

The actual information tasks and demands of modern business society are considered. The mathematical aspects of Bayesian Intelligent Technologies and experiences for economy application are shown.

The possibilities of methodological basis of RBA and RBA-technologies for actual social-economy tasks solving under conditions of dynamic situations, uncertainty and inaccuracy of the information, being of variety set of different types of data flows with risks definition, also the estimation of potentials for effective solving of these tasks and optimal resources distribution, have given.

CHAPTER 2

From fuzzy sets to soft estimates and synergetic artificial intelligence

V.B. Tarasov

***Abstract.** The paper presents an overview of fuzzy sets theory and reveals its role in the modeling of «im-in-un»-factors proper to human beliefs. Some differences between measurement and estimation are shown. The concepts of soft estimates and belief engineering are discussed. A linguistic-logical approach to taking various «im-in-un»s into account is suggested; fuzziness, contradiction and imprecision models are considered. The fundamentals of synergetic AI are discussed. The formal bilattices description is given in this context, their possibilities to become a unifying framework for non-standard fuzzy sets, as well as various cognitive and communicative problems are discussed.*

CHAPTER 3

Methods of constructing granular logical values and structures

V.B. Tarasov

***Abstract.** The paper is devoted to the problems of building granular logical values and structures. In the framework of universal logic approach some techniques and tools of logical granulation are considered: generalized truth-values and non-standard semantics, lattices, product lattices and bilattices, logical spaces and generalized logical matrices.*

CHAPTER 4

Modeling of opinions and assessments of intelligent agents: from four-valued systems of modalities to nonclassical measures and fuzzy allocations

V.B. Tarasov

***Abstract.** The paper deals with the modeling of basic cognitions related to intelligent agent activities, such as beliefs, judgments, estimates, norms, requirements, constraints, etc. The main emphasis is made on representing both truth and modal values in the framework of many-valued and fuzzy logics. The links between graded modalities, fuzzy distributions and non-classical measures are investigated. This approach based on crisp and fuzzy set functions serves as a unified framework to take into account partial truth, uncertainty level and inconsistency (relative contradiction) in agent's beliefs.*

CHAPTER 5

The theory of entropy potentials. State, development prospects and practical applications

V.L. Lasarev

***Abstract.** This Chapter presents information about the state of the theory of entropy potentials, taking into account recent developments in this direction. The prospects of its applications for solving a wide range of practical tasks in different situations and especially in the face of uncertainty. Considerable attention is paid to the application of the methods of this theory for monitoring, control and predicts the behavior of systems of different nature. The presentation is illustrated with examples.*

CHAPTER 6

Methodology of cognitive visualization of multidimensional data

V.L. Gorokhov, V.V. Vitkovskiy, Yu.V. Baryshev

***Abstract.** Principles and example of cognitive machine drawing for organization Decision Support Systems (DSS) are described. The cognitive machine drawing phenomenon essence is the special graphic representations, which create attractive images in the human operator brain. These images stimulate its descriptive impressions, closely related to the*

intuitive mechanisms of thinking. The cognitive effect lies in the fact that man perceives the moving projection as three-dimensional picture those characterized by multidimensional data properties in the multidimensional space. After the multidimensional data visual aspects study, it is appears the possibility to paint the interesting user separate objects or the groups of objects by standard machine drawing. Next user can return again to the image rotation procedure for checking the intuitive user's ideas about the clusters and the relationship in multidimensional data. It is possible to develop the cognitive machine drawing methods in combination with other information technologies. First of these are the packets of digital processing of images and multidimensional statistical analysis. The proposed method was based on the idea of possibility to assemble a cognitive image as object in hyperbolic space. In special sense it is perhaps to say that new kind of DSS – cognitive decision support systems (CDSS) are appears.

CHAPTER 7

Fuzzy methods of self-organization of information systems

M.S. Kupriyanov, **Л.Г. Логинская**

***Abstract.** In this section we discuss the method of qualitative scales re-grading for implementation of self-organization mechanisms and method of dynamic measurement of operability for self-repairing information systems.*

A fuzzy number is determined by the core - blur pair. Operations on fuzzy numbers are introduced. A procedure for the decomposition of a fuzzy value associated with a linguistic pair has been developed. Method of automatic qualitative scale re-grading was proposed on this basis.

Algorithms for operability evaluation of information systems and their operability management through self-organization are proposed.

CHAPTER 8

Methods for increasing the efficiency of genetic algorithms

M.S. Kupriyanov

***Abstract.** The fundamentals of the theory of Genetic Algorithms as well as the ways to improve efficiency of Genetic Algorithms are considered. Influence of selection, mutation and crossover operators has been studied.*

A Fuzzy Adaptive Crossover and Fuzzy Rule Base have been developed. Experimental efficiency studies of the algorithms with modified operators have been carried out.

CHAPTER 9

General concept and approach to the construction of distributed self-organizing information systems (ISIS)

M.S. Kupriyanov, **Ю.А. Шичкина**

***Abstract.** This section presents general principles for constructing of a distributed self-organizing information system (DSIS) as well as the mathematical model based on eight basic components of the system and the model of the system organization for Internet of Things. On the basis of the mathematical model of the DSIS, the corresponding structural-functional model and the system self-organization algorithm are presented. The Internet of things (IoT) system can be considered as an illustrative example of the DSIS. The problem arising during construction of the DSIS within the IoT concept is investigated at the end of the section. This is a task of data transmitting rules generation in Disruption Tolerant Network (DTN) on the basis of statistical information. The above solution of this problem illustrates the need to integrate technologies from different scientific fields, their interconnection, the depth of the problem solution and complexity of DSIS design.*

Научное издание

МЯГКИЕ ВЫЧИСЛЕНИЯ И ИЗМЕРЕНИЯ.

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ И МЕТОДЫ

Под редакцией проф. Прокопчиной Светланы Васильевны

Том I

Редактор О.В. Якупова
Компьютерная верстка С.Ю. Кирьянов
Корректор Н.А. Гежа

Подписано в печать 03.05.2017. Формат 60/90 1/16.
Бумага офсетная. Печать офсетная.
Объем 26,25 п.л. Тираж 1000 экз. Заказ _____

Издательский дом
«НАУЧНАЯ БИБЛИОТЕКА»
Телефон: 8 (495) 592-2998.
Адрес сайта: www.sciencelib.ru
E-mail: idnb11@yandex.ru, info@sciencelib.ru