Демидова Лилия Анатольевна

РАЗВИТИЕ МЕТОДОВ ТЕОРИИ НЕЧЁТКИХ МНОЖЕСТВ И ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ ДЛЯ ЗАДАЧ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЁННОСТИ

Специальность 05.13.01 — «Системный анализ, управление и обработка информации (технические системы)»

Автореферат диссертации на соискание ученой степени доктора технических наук

Работа выполнена в государственном образовательном учреждении высшего профессионального образования «Рязанский государственный радиотехнический университет»				
Научный консультант:	доктор технических наук, профессор, Заслуженный работник высшей школы Российской Федерации ПЫЛЬКИН Александр Николаевич			
Официальные оппоненты:	доктор технических наук, профессор Ковшов Евгений Евгеньевич			
	доктор технических наук, профессор Саксонов Евгений Александрович			
	доктор технических наук, профессор Скворцов Сергей Владимирович			
Ведущая организация:	МГТУ имени Н.Э. Баумана (г. Москва)			
	4 октября 2009 г. в 12 часов на заседании .01 в Рязанском государственном радиотех-90005, г. Рязань, ул. Гагарина, 59/1.			
С диссертацией можно ознаком венного радиотехнического универс	иться в библиотеке Рязанского государститета.			
Автореферат разослан «»	2009 г.			
Ученый секретарь диссертационного совета канд, техн. наук, доцент	В.Н. Пржегорлинский			

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Проблема принятия решений в условиях неопределенности занимает важное место в общей проблеме принятия решений. Успешное решение данной проблемы в настоящее время невозможно без применения новых информационных технологий, составной частью которых являются интеллектуальные средства обработки информации. Понятие «неопределенность» трактуется довольно неоднозначно, его смысл зависит от характера решаемой прикладной задачи. Для описания неопределенности современная теория принятия решений широко применяет, в частности, аппарат теории нечетких множеств (ТНМ), основоположником которой является Л.А. Заде (1965 г.). Проблема неопределенности присуща всем сложным системам. К таким сложным системам относятся: экономические и социально-экономические системы – системы прогнозирования показателей занятости населения в экономике страны; системы анализа инвестиционных решений и результатов внешнеторговой деятельности регионов и др.; технические системы – системы контроля, диагностики, классификации, кластеризации состояний объектов различного происхождения и др.

Традиционный подход к проблеме принятия решений основан на использовании классических методов многокритериального анализа (А.А. Амосов, А. Вальд, Р. Кини, Р. Клемен (R. Clemen), О.И. Ларичев, В.Д. Ногин, А.И. Орлов, Т. Л. Саати, Дж. К. Смит (J.Q. Smith), А.Н. Тихонов, С. Ханссон (S. Hansson)) и предполагает разработку и создание сложных, зачастую многоуровневых, систем поддержки принятия решений, базирующихся на математических моделях, обеспечивающих учет большого количества параметров и критериев, и характеризуется значительными вычислительными затратами и высокой стоимостью разработки.

Задачи поддержки принятия решений в условиях неопределенности представляют собой слабоструктурированные или неструктурированные задачи. Слабоструктурированные задачи содержат неизвестные или неизмеряемые компоненты, то есть количественно неоцениваемые компоненты. Такие задачи характеризуются отсутствием методов решения на основе непосредственных преобразований данных, а постановки задач предполагают принятие решений в условиях неполной информации. Неструктурированные задачи содержат неформализуемые процедуры, базирующиеся на неструктурированной информации, которая определяется высокой степенью неопределенности. Применение ТНМ и её приложений позволяет построить формальные схемы решения задач, характеризующихся той или иной степенью неопределенности, которая может быть обусловлена неполнотой, внутренней противоречивостью, неоднозначностью и размытостью исходных данных, представляющих собой приближенные количественные или качественные оценки параметров объектов. Эта неопределенность является систематической, так как обусловлена сложностью задач, дефицитом информации, лимитом времени на принятие решений. особенностями восприятия и т.п.

Неполнота и неточность информации могут заключаться: в принципиальной невозможности полного сбора и учета информации об анализируемом объекте или процессе; в некоторой недостоверности и недостаточности исходной информации об анализируемом объекте или процессе; в возможности проявления таких свойств анализируемого объекта или процесса, существование которых не предполагалось. Кроме того, неточность, неполнота и неопределенность исходных данных могут быть вызваны недостаточными знаниями экспертов специфики конкретной прикладной задачи. Следовательно, можно говорить и о наличии «субъективного» человеческого фактора в задачах поддержки принятия решений в условиях неопределенности. В этом случае исходные данные, представленные качественными оценками объектов, могут быть не только несовпадающими, но и противоречивыми, что значительно усложняет решение задачи поддержки принятия решений. Процессы, представленные временными рядами (ВР), в реальных условиях также характеризуются неполнотой информации, которая может быть принципиальной из-за неповторимости (невоспроизводимости) явления.

Таким образом, использование классических «жестких» алгоритмов моделирования сложных систем, предполагающих регламентированную постановку задачи и пошаговый процесс получения результатов, при решении различных задач поддержки принятия решений может оказаться малоэффективным, так как предполагает обработку точных, полных и непротиворечивых численных исходных данных, что может не соответствовать высокому уровню неопределенности задачи.

Отказ от традиционных требований точности измерений, которая была необходима при математическом анализе четко определенных систем и процессов, и применение ТНМ совместно с методами алгебры логики обеспечивают решение проблемы принятия решений в условиях неопределенности. Принципиальной особенностью задач поддержки принятия решений в условиях неопределенности является необходимость учета того факта, что измерения входных и выходных данных выполняются на уровне «мягких измерений». Использование понятия «лингвистическая переменная» позволяет адекватно отразить приблизительное словесное описание некоторых параметров и состояний объекта или процесса в тех случаях, когда точное описание либо отсутствует, либо является слишком сложным, либо требует больших временных и финансовых затрат.

При разработке алгоритмов и методов ТНМ охватывается широкий круг математических и прикладных проблем, в решение которых значительный вклад внесли российские и зарубежные ученые: А.Н. Аверкин, А.В. Алексеев, Р.А. Алиев, А.Е. Алтунин, К. Асаи, И.З. Батыршин, Р. Беллман, Л.С. Бернштейн, А.Н. Борисов, В.В. Борисов, Л.А. Заде, С.Я. Коровин, А. Кофман, О.А. Крумберг, А.В. Леоненков, Н.Г. Малышев, А.Н. Мелехов, К. Негойце, С.А. Орловский, Д.А. Поспелов, Р. Ягер, Т.Л. Саати, М.В. Семухин, В.Б. Силов, А. Тверски, Т. Тэрано, С.Д. Штовба и др. Вопросы реализации алгоритмов нечеткого вывода рассматриваются в работах X. Ларсена (H. Larsen), Е. Мамдани (E. Mamdani), М. Cyreno (M. Sugeno), Т. Такаги (Т. Takagi), Й. Цукамото (Ý. Tsukamoto). В работах Дж.К. Беждека (J.C. Bezdek), Дж.К. Данна (J.C. Dunn), Р.Н. Дейва (R.N. Dave), Дж.М. Келлера (J.M. Keller), Р. Кришнапурама (R. Krishnapuram), Я. Охаши (Y. Ohashi) предложены алгоритмы нечеткой кластеризации (алгоритм нечетких с -средних и его модификации). Значительное количество работ (Я. Батистакис (Y. Batistakis), Г. Бени (G. Beni), Х. Галда (H. Galda), И. Гаф (I. Gath), А. Гева (A.B. Geva), Д. Густафсон (D. Gustafson), В. Кессел (W. Kessel), M. Сугено, M. Халкиди (M. Halkidi), C. Če (X. Xei), Ю. Фукуяма (Y. Fukuyama)) посвящено разработке и исследованию показателей качества кластеризации с использованием алгоритмов нечеткой кластеризации. Проблема разработки эффективных моделей прогнозирования на основе нечетких временных рядов рассматривается в работах Л. Х. Ванга (L.H. Wang), Л.В. Ли (L.W. Lee), К. Сонга (Q. Song), С.М. Чена (S.M. Chen), Б.С. Чиссома (B.S. Chissom).

В настоящее время всё большее внимание уделяется разработке гибридных подходов к многокритериальному анализу сложных систем, основанных на «мягких» вычислениях и реализующих совместное применение различных методов искусственного интеллекта, позволяющих сформировать новую информационную технологию, важную роль в которой играют знания предметной области конкретной прикладной задачи (В.В. Борисов, В.В. Круглов, А.В. Кузьмин, Ю.Н. Минаев, М. Пилиньский, А.П. Ротштейн, Д. Ругковская, Л. Ругковский, А.А. Усков, О.Ю. Филимонова, А.С. Федулов, Н.Г. Ярушкина). При этом этапы решения задачи и результаты определяются текущим состоянием базы знаний, а не каким-либо «жестким» алгоритмом моделирования. К таким методам, в первую очередь, относятся методы, основанные на применении искусственных нейронных сетей, использовании инструментария ТНМ и теории генетических алгоритмов (ГА).

Одним из современных бионических принципов решения широкого класса прикладных задач, которые трудноразрешимы классическими методами, особенно в области NP -полных задач оптимизации, является применение ГА – адаптивных методов поиска, реализующих эволюционные вычисления, основанные на генетических процессах биологических организмов. Общие принципы ГА были сформулированы Д.Х. Холландом

(1975 г.) и описаны в работах: Д.И. Батищева, Л.А. Гладкова, Д.И. Голдберга, В.В. Емельянова, Е.Е. Ковшова, В.В. Курейчика, В.М. Курейчика и др.

В последние годы наряду с обычными нечеткими множествами (нечеткими множествами первого типа — НМТ1) большее применение в решении различных прикладных задач находят интервальные нечеткие множества второго типа (ИНМТ2), использование которых, однако, сопровождается увеличением вычислительной сложности алгоритмов. Существенный рост количества прикладных задач, решаемых с использованием ИНМТ2, связан с публикацией в 2001 году основополагающей статьи Н. Карника (N. Karnik) и Дж. М. Менделя (J.М. Mendel), в которой предложен итерационный алгоритм вычисления центроида ИНМТ2, реализующий операции «понижения» типа и дефаззификации и значительно снижающий сложность вычисления центроида ИНМТ2. Тем не менее, использование ИНМТ2 целесообразно, если ожидается существенное улучшение результатов (например, повышение точности прогнозирования, улучшение качества кластеризации). Вопросы разработки систем нечеткого вывода на основе ИНМТ2 отражены в работах О. Кастилло (О. Castillo), С. Коупланда (S. Coupland), П. Мелина (Р. Melin), Дж.М. Менделя (J.М. Mendel), основные принципы алгоритмов кластеризации на основе ИНМТ2 изложены в работе Ф.Ч.-Х. Рхи (F.С.-Н. Rhee) и Ч. Хванга (С. Hwang) (2007 г.).

Совместное использование инструментария ТНМ и теории мультимножеств, основные положения которой изложены в работах А.Б. Петровского, обеспечивает принятие адекватных и обоснованных решений в случае наличия несовпадающих и даже противоречивых исходных данных.

Анализ известных методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности, основанных на использовании инструментария ТНМ, показывает, что довольно часто они не обеспечивают получение адекватных решений ввиду недостаточно обоснованного выбора параметров моделирования, а поиск эффективных решений сопровождается значительными временными затратами из-за необходимости выполнения многократных реализаций используемых методов, моделей и алгоритмов с целью выбора оптимальных параметров.

Актуальность настоящей работы определяется необходимостью разработки эффективных методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности, обеспечивающих высокую обоснованность и адекватность принимаемых решений при низких временных затратах. Использование различных модификаций ГА позволяет решить проблему выбора оптимальных параметров методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности при приемлемых временных затратах. Комплексное использование инструментария ТНМ, нечеткой логики, теории мультимножеств и генетических алгоритмов позволяет создать качественно новые программные средства, существенно расширяющие перечень рассматриваемых задач поддержки принятия решений в условиях неопределенности и обеспечивающие повышение точности, адекватности и объективности (а следовательно, и эффективности) принятия решений в условиях неопределенности.

<u>Объект исследования.</u> Объектом диссертационного исследования являются методы, модели и алгоритмы поддержки принятия решений в условиях неопределенности, а также их практические реализации в системах поддержки принятия решений.

Предмет исследования. Предметом исследования являются:

- 1. Модели прогнозирования процессов с преобладающей детерминированной составляющей, представленных временными рядами с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса.
 - 2. Модели поддержки принятия решений с использованием систем нечеткого вывода.
- Методы и алгоритмы упорядочения, классификации и кластеризации многомерных объектов при ярко выраженном наличии несовпадающих, в том числе противоречивых, оценок экспертов с использованием подхода теории мультимножеств.

4. Методы и алгоритмы кластеризации многомерных объектов в условиях неполноты и нечеткости исходной информации.

<u>Цель диссертационной работы</u> состоит в разработке эффективных методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности на основе комплексного использования инструментария теории нечётких множеств и генетических алгоритмов, позволяющего устранить недостатки существующих аналогов, обеспечивая при этом высокую обоснованность и адекватность принимаемых решений.

Для достижения поставленной цели решались следующие задачи.

- 1. Разработка моделей краткосрочного прогнозирования процессов с преобладающей детерминированной составляющей, представленных временными рядами с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса, на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа с использованием генетических алгоригмов.
- Разработка моделей поддержки принятия решений с использованием систем нечеткого вывода на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа и генетических алгоритмов «тонкой» настройки параметров систем нечеткого вывода.
- 3. Разработка методов упорядочения, классификации и кластеризации объектов при наличии несовпадающих, в том числе противоречивых, исходных данных с использованием мультимножеств, нечеткого метода Дельфы, схемы Беллмана Заде и систем нечеткого вывода.
- 4. Разработка методов и алгоритмов кластеризации объектов и соответствующих показателей качества кластеризации с использованием модификаций алгоритма нечетких с -средних на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа и генетических алгоритмов.
- Решение ряда актуальных прикладных задач поддержки принятия решений в условиях неопределенности с использованием разработанных методов, моделей и алгоритмов.

Методы исследований. Методы исследований, используемые в данной работе, объединяются на основе системного подхода к решению поставленных задач. Теоретические исследования выполнены с использованием методов теории нечетких множеств, нечеткой логики, теории генетических алгоритмов, теории мультимножеств, теории сложности, теории вероятностей, математической статистики, математического и системного анализа, аналитической геометрии; экспериментальные исследования выполнены с привлечением методов математического и имитационного моделирования, технологий модульного и объектно-ориентированного программирования.

<u>Научная новизна.</u> В рамках диссертационной работы были получены следующие основные результаты, обладающие научной новизной.

- 1. Разработаны модели краткосрочного прогнозирования процессов с преобладающей детерминированной составляющей, представленных временными рядами с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса, на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа с использованием генетических алгоритмов, обеспечивающие повышение точности прогнозирования за счет выбора оптимальных параметров моделей прогнозирования.
- 2. Разработаны генетические алгоритмы «тонкой» настройки параметров систем нечеткого вывода на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа с использованием обучающих выборок, обеспечивающие повышение объективности и адекватности принимаемых решений с применением моделей поддержки принятия решений в условиях неопределенности.
- 3. Разработаны методы упорядочения, классификации и кластеризации объектов при наличии несовпадающих, в том числе противоречивых, исходных данных с использова-

нием мультимножеств, нечеткого метода Дельфы, схемы Беллмана — Заде и систем нечеткого вывода, позволяющие учесть в процессе принятия решения все, в том числе противоречивые, оценки объектов без использования дополнительных преобразований типа усреднения и смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных. Предложено при формировании обобщающих правил классификации объектов, представленных мультимножествами, для выполнения предварительной сортировки объектов применять индивидуальные системы нечеткого вывода, параметры которых настроены с использованием генетических алгоритмов.

- 4. Предложен метод оценивания объектов с вычислением центроидов интервальных нечетких множеств второго типа на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы, позволяющий дифференцировать объекты, имеющие одинаковые центроиды нечетких множеств первого типа на основе этих же экспертных оценок.
- 5. Разработаны методы кластеризации объектов с использованием модификаций алгоритма нечетких c-средних на основе нечетких множеств первого типа и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры подобной плотности и подобного объема, с минимальными временными затратами.
- 6. Разработаны методы кластеризации объектов с использованием модификаций алгоритма нечетких c -средних на основе интервальных нечетких множеств второго типа и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры существенно разной плотности или существенно разного объема, с минимальными временными затратами.
- 7. Разработаны показатели качества кластеризации, позволяющие получить адекватные результаты кластеризации множества объектов, содержащего кластеры гиперсферической или гиперэллипсоидной формы, с использованием методов кластеризации на основе интервальных нечетких множеств второго типа.

<u>Теоретическая значимость работы</u> заключается в обобщении теории и развитии методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности с использованием бионических принципов решения прикладных задач.

<u>Практическая ценность работы.</u> Практическая ценность работы состоит в том, что разработанные методы, модели и алгоритмы поддержки принятия решений позволяют:

- обеспечить высокую обоснованность и адекватность принятия решения в условиях неопределенности и неточности исходной информации, в том числе при несовпадающих (противоречивых) оценках экспертов;
- обоснованно использовать формализацию опыта экспертов, который зачастую является единственной наиболее достоверной информацией при решении многих задач поддержки принятия решений;
- минимизировать временные и финансовые затраты, связанные как с необходимостью сбора и учета точных и полных исходных данных (что может быть принципиально невозможным), так и с разработкой сложных классических математических моделей или необходимостью многократной реализации классических методов, моделей и алгоритмов с целью выбора соответствующих оптимальных параметров, обеспечивающих принятие адекватных решений.

<u>Достоверность</u> научных положений, теоретических выводов и практических результатов диссертационной работы подтверждается:

- корректным использованием понятий и выводов теории нечетких множеств и нечеткой логики, теории мультимножеств и теории генетических алгоритмов;
- результатами математического и имитационного моделирования разработанных методов, моделей и алгоритмов;

- разработкой действующих программных средств, подтвержденных свидетельствами об официальной регистрации;
- апробацией предложенных разработанных методов, моделей и алгоритмов на конкретных прикладных задачах;
 - наличием актов внедрения результатов диссертационной работы.

Реализация и внедрение результатов диссертационной работы

Исследования по тематике диссертационной работы велись: в рамках Государственной научной стипендии с 01.04.1997 по 31.03.2000 (Указ Президента Российской Федерации «О мерах по материальной поддержке ученых России» № 1372 от 16.09.1993; Постановление Президиума Российской академии наук № 66 от 15.04.1997); в рамках 5 госбоджетных НИР: НИР 24-03Г «Разработка и исследование математических моделей и алгоритмов прогнозирования социально-экономических процессов» (2003 г.); НИР 1-03Г «Разработка и исследование математических моделей представления и обработки информации в вычислительных системах» (2003-2005 гг.); НИР 4-04Г «Разработка методов прогнозирования социально-экономических процессов на основе процедур квалификации наблюдений, самоорганизации формальных описаний мониторинга данных» (2004-2006 гг.); НИР 9-07Г «Разработка математических моделей, методов и алгоритмов обработки больших потоков информации в сложно организованных вычислительных структурах» (2007-2008 гг.); НИР 10-09Г «Методы и алгоритмы идентификации и прогнозирования состояния больших систем в условиях неопределенности на основе нечеткой логики, генетических алгоритмов и мультимножеств» (2009 г.).

Разработанные методы, модели и алгоритмы, а также реализующие их пакеты прикладных программ (ППП) внедрены в организациях г. Рязани.

- 1. ППП «Модели прогнозирования на основе нечетких множеств первого и второго типов» показал эффективность при разработке краткосрочных прогнозов показателей занятости населения РФ в экономике страны и при выполнении метеопрогнозов в городе Камбарка Удмуртской Республики. ППП внедрен и используется в ООО «НЭК Квандо-Терминал» для прогнозирования показателей, характеризующих внешнеторговую деятельность: «таможенные платежи», «индекс таможенной стоимости», «экспорт товаров», «импорт товаров», «количество Государственных таможенных деклараций, оформленных в режиме экспорта и импорта для свободного обращения» и др. Экспериментальные исследования показали целесообразность использования ППП для краткосрочного прогнозирования (на 1 шаг вперед) процессов с преобладающей детерминированной составляющей, представленных ВР с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса.
- 2. В управлении топливно-энергетического комплекса и жилищно-коммунального хозяйства Рязанской области и в МУП «Рязанское муниципальное предприятие тепловых сетей» успешно внедрен и используется ППП «Нечеткие городские инженерные коммуникации (ГИК)», предназначенный для решения задач диагностики состояния водопроводных сетей, теплосетей и арматуры в штатных и аварийных ситуациях при необходимости учёта территориальной распределённости объектов. Опытная эксплуатация подтвердила работоспособность и показала высокие характеристики надежности ППП.
- 3. В Рязанской торгово-промышленной палате и в ЗАО «ПРО-САМ» внедрен и используется ППП «Упорядочение и классификация инвестиционных проектов на основе мультимножеств», предназначенный для многокритериального анализа инвестиционных проектов (ИП) при ярко выраженном наличии несовпадающих, в том числе противоречивых, оценок экспертов.
- 4. ППП «Настройка систем нечеткого вывода классификации способов несанкционированного отбора электроэнергии с использованием генетических алгоритмов» используется в Филиале «Рязаньэнерго» ОАО «МРСК Центра и Приволжья» при решении задачи классификации способов несанкционированного отбора электроэнергии для вы-

полнения настройки параметров систем нечеткого вывода абонентов электросети с использованием их индивидуальных моделей энергопотребления при наличии графического и/или аналитического представления данных обучающей выборки.

- 5. ППП «Оценка обобщенного мнения экспертов на основе нечеткого метода Дельфы и центроидов нечетких множеств первого и второго типов» и ППП «Методы кластеризации объектов на основе нечетких множеств первого и второго типов» внедрены и используются в деятельности ООО «Независимый центр оценки и экспертиз» при решении задачи оценки технического состояния зданий и сооружений как задачи кластеризации при нечётком определении состояний многомерных объектов и задачи оценки недвижимости как задачи оценки потенциального качества изменяемых многомерных объектов при малых объёмах априорной информации соответственно. Опытная эксплуатация показала целесообразность использования первого ППП при формировании обобщенного мнения экспертов с использованием центроидов НМТ1, а в сложных случаях центроидов ИНМТ2, применение которых позволяет дифференцировать объекты, имеющие равные центроиды НМТ1 для оценок по некоторым критериям. Второй ППП показал высокую эффективность при решении задач многокритериальной кластеризации совокупностей объектов, содержащих как кластеры подобной плотности и подобного объема, так и кластеры существенно разного объема.
- 6. Результаты полученных в диссертации теоретических, прикладных и экспериментальных исследований используются в учебном процессе Рязанского государственного радиотехнического университета при обучении студентов специальности 230105 «Программное обеспечение вычислительной техники и автоматизированных систем» в курсе «Проектирование искусственного интеллекта», специальности 080801 «Прикладная информатика в экономике» в курсе «Информационные технологии» и «Элементы теории нечетких множеств», специальности 010503 «Математическое обеспечение и администрирование информационных систем» в курсе «Элементы теории нечетких множеств».

Использование результатов диссертационной работы на практике подтверждено соответствующими актами о внедрении.

Получено 6 свидетельств ФГУ «Федеральный институт промышленной собственности Федеральной службы по интеллектуальной собственности, патентам и товарным знакам» (ФГУ ФИПС – РОСПАТЕНТ) и 9 свидетельств ОФАП об официальной регистрации программ для ЭВМ.

Основные положения, выносимые на защиту

- 1. Модели краткосрочного прогнозирования процессов с преобладающей детерминированной составляющей, представленных временными рядами с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса, на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа с использованием генетических алгоритмов, обеспечивающие повышение точности прогнозирования за счет выбора оптимальных параметров моделей прогнозирования.
- 2. Генетические алгоритмы «тонкой» настройки параметров систем нечеткого вывода на основе нечетких множеств первого типа и интервальных нечетких множеств второго типа с использованием обучающих выборок, обеспечивающие повышение объективности и адекватности принимаемых решений с применением моделей поддержки принятия решений в условиях неопределенности.
- 3. Методы упорядочения, классификации и кластеризации объектов при наличии несовпадающих, в том числе противоречивых, исходных данных с использованием мультимножеств, нечеткого метода Дельфы, схемы Беллмана — Заде и систем нечеткого вывода, позволяющие учесть в процессе принятия решения все, в том числе противоречивые, оценки объектов без использования дополнительных преобразований типа усреднения и смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных.

- 4. Метод оценивания объектов с вычислением центроидов интервальных нечетких множеств второго типа на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы, позволяющий дифференцировать объекты, имеющие одинаковые центроиды нечетких множеств первого типа на основе этих же экспертных оценок.
- 5. Методы кластеризации объектов с использованием модификаций алгоритма нечетких с-средних на основе нечетких множеств первого типа и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры подобной плотности и подобного объема, с минимальными временными затратами.
- 6. Методы кластеризации объектов с использованием модификаций алгоритма нечетких c -средних на основе интервальных нечетких множеств второго типа и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры существенно разной плотности или существенно разного объема, с минимальными временными затратами.
- 7. Показатели качества кластеризации, позволяющие получить адекватные результаты кластеризации множества объектов, содержащего кластеры гиперсферической или гиперэллипсоидной формы, с использованием методов кластеризации на основе интервальных нечетких множеств второго типа.
- Результаты решения ряда актуальных прикладных задач поддержки принятия решений в условиях неопределенности с использованием разработанных методов, моделей и алгоритмов.

Апробация работы. Основные научные положения диссертационной работы докладывались и обсуждались на научных конференциях: 11-й Всероссийской межвузовской научно-технической конференции студентов и аспирантов «Микроэлектроника и информатика» (Москва, 2004); VIII Международной научно-практической конференции «Системный анализ в проектировании и управлении» (Санкт-Петербург, 2004); 7-й Всероссийской с Международным участием научной конференции молодых ученых и аспирантов «Новые информационные технологии. Разработка и аспекты применения» (Таганрог, 2004); 11-й Международной научно-технической конференции студентов и аспирантов «Радиоэлектроника, электротехника и энергетика» (Москва, 2005): Международном форуме по проблемам науки, техники и образования (Москва, 2005); 1-й Всероссийской конференции изыскательских организаций «Перспективы развития инженерных изысканий в строительстве в Российской Федерации» (Москва, 2006); 33-й Всероссийской научно-технической конференции «Сети, системы связи и телекоммуникации. Деятельность ВУЗа при переходе на Федеральный государственный образовательный стандарт 3-го поколения» (Рязань, 2008); 12, 13, 14 и 15-й Международных научнотехнических конференциях «Проблемы передачи и обработки информации в сетях и системах телекоммуникаций» (Рязань, 2004, 2004, 2005, 2008); 10 и 14-й Международных открытых научных конференциях «Современные проблемы информатизации в технике и технологиях» (Воронеж, 2005, 2009); IV и VI Всероссийских научно-технических конференциях «Искусственный интеллект в XXI веке. Решения в условиях неопределенности» (Пенза, 2006, 2008); VII и VIII Всероссийских научно-технических конференциях «Проблемы информатики в образовании, управлении, экономике и технике» (Пенза, 2007, 2008); X и XIII Всероссийских научно-технических конференциях студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях и образовании» (Рязань, 2005, 2008); Международных научно-практических конференциях «Управление созданием и развитием систем, сетей и устройств телекоммуникаций» (Санкт-Петербург, 2008).

Публикации. По теме диссертации опубликовано 77 печатных работ (13 — без соавторов), в том числе: 17 статей в изданиях, входящих в перечень ВАК для докторских диссертаций; 2 монографии; 7 статей в научно-технических журналах; 12 статей в межвузовских сборниках; 24 доклада на Международных и Всероссийских конференциях; 15 свидетельств о регистрации программ: 6 свидетельств о регистрации программ для ЭВМ в ФГУ «Федеральный институт промышленной собственности Федеральной службы по интеллектуальной собственности, патентам и товарным знакам» (ФГУ ФИПС — РОСПАТЕНТ); 9 свидетельств об официальной регистрации программ в Отраслевом фонде алгоритмов и программ (ОФАП).

Личный вклад автора в получение результатов, изложенных в диссертации. Все результаты диссертационной работы, в том числе постановка задач, разработка и исследование защищаемых методов, моделей и алгоритмов, основные научные результаты, выводы и рекомендации принадлежат лично автору. Программные средства, реализующие разработанные методы, модели и алгоритмы, разработаны под руководством и при непосредственном участии автора. Работы, выполненные в соавторстве, посвящены общей постановке проблемы, концепции ее решения, предложенной автором, конкретизации разработанных методов, моделей и алгоритмов для ряда актуальных прикладных задач, разработке отдельных программных средств.

Структура и объем диссертации. Диссертационная работа состоит из введения, шести глав, списка литературы, заключения, изложенных на 532 страницах (включая 74 рисунка и 7 таблиц), и отдельного тома, содержащего четыре приложения на 456 страницах (включая 265 рисунков и 181 таблицу). Список литературы состоит из 400 наименований.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы, сформулированы цель и задачи диссертационной работы, определены направления исследования. Показаны научная новизна и практическая ценность диссертационной работы.

В главе 1 «Обзор и анализ методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности» показано, что проблема поддержки принятия решений в условиях неопределенности относится к классу слабо структурированных задач и их решение зависит от типа объекта, характера неопределенностей и возможностей моделирования объекта. Характерной особенностью таких объектов являются невозможность (или ограниченная возможность) моделирования в реальном времени и необходимость выработки решения на основе моделирования.

Проведен анализ теорий неопределенностей, показавший преимущества использования ТНМ для решения задач поддержки принятия решений в условиях неопределенности. Многие задачи, связанные с проблемой поддержки принятия решений в условиях неопределенности, относятся к классу NP -полных задач, решение которых возможно с использованием вложенных методов «мягких вычислений», реализующих бионические принципы решения задач: ГА, искусственных нейронных сетей и искусственных иммунных сетей. Показано, что ГА являются мощным инструментом решения широкого спектра прикладных задач, включая NP -полные задачи оптимизации. Определены области применения нечеткого моделирования: недостаточные (неопределенные) знания об исследуемой системе, адекватная обработка неопределенной информации, «прозрачное» моделирование и идентификация систем. Дана классификация нечетких моделей: для описания системы; задания параметров; задания входов, выходов и состояний системы. Определены области эффективного применения классических, нейросетевых и нечетких моделей. Классические модели следует использовать при относительно невысокой сложности системы и наличии достаточно полной информации о системе; нейросетевые модели следует применять при отсутствии информации о системе или при высокой сложности системы; применение нечетких моделей целесообразно при средней сложности системы и частичном наличии информации (неполной, неточной, нечеткой) о системе. При наличии обучающей выборки нечеткие модели позволяют аппроксимировать измеренные данные с требуемой точностью, что позволяет считать их универсальными аппроксиматорами.

Даны определения НМТ1 и ИНМТ2. НМТ1 представляется в виде $A = \{(x, u_A(x)) | x \in X, u_A(x) \in [0,1]\}$, где $u_A(x)$ — функция принадлежности (ФП). При описании неопределенности с помощью ИНМТ2 первичная ФП J_x определяется как интервальная ФП с вторичными степенями первичных ФП, равными 1. ИНМТ2 \widetilde{A} представляется как $\widetilde{A} = \{((x,u),\mu_{\widetilde{A}}(x,u)) | \forall x \in X, \forall u \in J_x \subseteq [0,1], \mu_{\widetilde{A}}(x,u) = 1\}$. Использование ИНМТ2 целесообразно только в том случае, когда ожидается существенное улучшение результатов (например, повышение точности прогнозирования, улучшение качества кластеризации, улучшение результатов аппроксимации и т.п.) ввиду увеличения сложности вычислений.

Сформулированы принципы анализа многокритериальных задач принятия решений, в том числе в условиях неопределенности. Показана симметричность нечетких целей и ограничений относительно решения в схеме Беллмана — Заде, так как решение представляется как слияние нечетких целей и ограничений. Определены принципы упорядочения, классификации и кластеризации объектов в задачах многокритериального анализа. Показано, что в случае группового экспертного оценивания объектов в качестве математической модели для представления многокритериальных объектов следует использовать мультимножество, позволяющее учесть все, в том числе несовпадающие и противоречивые, оценки объектов по критериям. Выполнен анализ целесообразности разработки и использования алгоритмов кластеризации на основе нечетких множеств. Выявлена проблема кластеризации множеств объектов, содержащих кластеры существенно разной мощности, существенно разной плотности и т.п.

Показано, что использование нечетких временных рядов (НВР) является обоснованным в задачах анализа и моделирования динамики коротких ВР, для которых проблема анализа еще не получила эффективного решения классическими методами. Особое внимание должно уделяться поиску оптимальных параметров модели прогнозирования (в смысле некоторого критерия), что может быть реализовано с использованием ГА.

Сформулированы актуальные задачи поддержки принятия решений в условиях неопределенности, свидетельствующие о том, что подходы к принятию решений в условиях неопределенности на основе бионических принципов являются универсальными, однако требуют тщательной проработки и адаптации к конкретной прикладной задаче.

Глава 2 «Модели прогнозирования на основе нечетких временных рядов и генетических алгоритмов» посвящена исследованию моделей краткосрочного прогнозирования (на 1 шаг вперед) на основе НВР и разработке моделей прогнозирования на основе НМТ1 и ИНМТ2 с использованием модификаций ГА, реализующих поиск оптимальных параметров моделей прогнозирования, обеспечивающих минимизацию средней относительной ошибки прогнозирования *AFER* (average forecasting error rate), по значению которой оценивается качество модели. В качестве базовой модели рассматривается модель прогнозирования на основе дискретных нечетких множеств первого типа (ДНМТ1), предложенная С.М. Ченом¹ (2002 г.).

Пусть d(t) (t=0,1,2,...,m) — BP на основе реальных значений некоторого показателя (фактора). BP на основе значений приращений показателя представляется как $\Delta d(t) = d(t) - d(t-1)$ (t=0,1,2,...,m). Значения BP целесообразно представлять с помощью значений приращений показателя, что обеспечивает получение меньших значений AFER. Значение HBP представляется с помощью ДНМТ1 A, определенного на универсуме X:

10

¹ Chen S.M. Forecasting enrollments based on high-order fuzzy time series // Cybernetic Systems, 2002. – V. 33. – № 1. – P. 1-16.

 $A = u_A(x_1)/x_1 + u_A(x_2)/x_2 + ... + u_A(x_n)/x_n$, где $u_A(x)$ — функция принадлежности ДНМТ1 A, $u_A(x): X \to [0, 1]$, $u_A(x_r)$ ($r = \overline{1, n}$) определяет степень принадлежности элемента x_r .

Пусть Y(t) (t = ..., 0,1,2,...) — универсум на множестве действительных чисел. Зададим на универсуме Y(t) нечеткие множества (HM) $u_i(t)$ (i = 1,2,...). Пусть U(t) — набор функций $u_i(t)$, определенных на универсуме Y(t); U(t) называется HBP на универсуме Y(t). Если U(t+1) зависит от U(t),..., U(t-k), U(t-k+1), то нечеткая логическая зависимость (НЛЗ) представляется как $U(t-k+1), U(t-k),..., U(t) \to U(t+1)$ и называется однофакторной моделью прогнозирования t —го порядка на основе HBP.

Модель прогнозирования на основе НВР предполагает: определение универсума и интервалов принадлежности для значений ВР; определение лингвистических термов; фаззификацию значений ВР; формирование групп нечетких логических зависимостей (ГНЛЗ); прогнозирование; дефаззификацию результатов прогнозирования.

Рассмотрим модель прогнозирования на основе ДНМТ1. Представим t -е и (t+1) -е значения BP, соответствующие некоторым отсчётам времени, как НМ A_j и A_l . НЛЗ может быть записана как $A_j \to A_l$. Под отсчётом времени для BP будут пониматься день, месяц, квартал и т.п. Универсум X для значений приращений показателя определяется как $X = [D_{min} - D_1, D_{max} + D_2]$, где D_{min} и D_{max} — минимальное и максимальное значения приращения показателя на основе BP $\Delta d(t)$ (t=1,2,...,m) соответственно ($D_{min} = \min_{t=1,m} (\Delta d(t))$), $D_{max} = \max_{t=1,m} (\Delta d(t))$), а D_l и D_2 — положительные действительные числа, использование которых позволяет разбить универсум X на n интервалов равной длины: x_1 , x_2 ,..., x_n . Лингвистические термы A_r ($r=\overline{1,n}$), представленные ДНМТ1 для значений приращений показателя, могут быть записаны в виде:

Фаззификация данных выполняется следующим образом. Если значение приращения показателя принадлежит интервалу x_1 , то нечеткое значение имеет вид: $W_1=1/A_1+0.5/A_2$; если значение приращения показателя принадлежит интервалу u_r , то нечеткое значение имеет вид: $W_r=0.5/A_{r-1}+1/A_r+0.5/A_{r+1}$, $r=\overline{2,n-1}$; если значение приращения показателя принадлежит интервалу x_n , то нечеткое значение имеет вид: $W_n=0.5/A_{n-1}+1/A_n$. Пусть W_j и W_l — нечеткие значения приращений показателя для t-го и (t+1)-го отсчетов времени, определяющие НЛЗ вида: $W_j\to W_l$. Для всех известных значений ВР находятся НЛЗ. Затем НЛЗ, имеющие одинаковую левую часть, объединяются в одну ГНЛЗ. Так, если были сформированы НЛЗ первого порядка: $W_j\to W_{l_1}$, $W_j\to W_{l_2}$, то они объединяются в группу:

$$W_j \to W_{l_1}, W_{l_2}, ..., W_{l_g}$$
 (3)

Аналогичным образом формируются ГНЛЗ высокого порядка. Прогнозирование может выполняться как без учета повторов НМ в правых частях ГНЛЗ, так и с учетом повторов НМ.

В первом случае результирующее ДНМТ1 для прогнозируемого значения BP для (t+1) -го отсчета времени находится как объединение ДНМТ1, входящих в правую часть ГНЛЗ для (t+1) -го отсчета времени, с ФП вида:

$$u_{W}(x_{r}) = \max(u_{W_{l_{1}}}(x_{r}), u_{W_{l_{2}}}(x_{r}), \dots, u_{W_{l_{g}}}(x_{r})) \quad (r = 1, n).$$

$$(4)$$

Искомое значение прогнозируемой величины f(t+1) для (t+1) -го отсчета времени находится как сумма значения d(t), определяемого с помощью BP на основе реальных значений показателя для t -го отсчета времени, и дефаззифицированного значения приращения показателя y(t+1) для (t+1) -го отсчета времени: f(t+1) = d(t) + y(t+1). Дефаззифицированное значение приращения показателя y(t+1) для (t+1) -го отсчета времени определяется по методу центра тяжести для одноточечных множеств:

$$y(t+1) = \sum_{r=1}^{n} \left(w_r(t+1) \cdot \chi_r \right) / \sum_{r=1}^{n} w_r(t+1) ,$$
 (5)

где n — количество интервалов x_r ($r=\overline{1,n}$); χ_r — средняя точка r -го интервала x_r ; $w_r(t+1)$ — значение степени принадлежности для r -го интервала результирующего ДНМТ1 для (t+1) -го отсчета времени в правой части ГНЛЗ вида (3) с ФП в соответствии с формулой (4) (при этом $w_r(t+1)$ могут принимать значения 0, 1 или 0,5, что следует из соотношений вида (2)).

Учет повторяющихся ДНМТ1 в правых частях ГНЛЗ позволяет уменьшить значение AFER. Если при формировании ГНЛЗ выявлены повторяющиеся ДНМТ1 в правых частях групп, то значение приращения показателя y(t+1) для (t+1) -го отсчета времени может быть вычислено по формуле:

$$y(t+1) = \sum_{r=1}^{n} \left(v_r(t+1) \cdot \tau_r \right) / \sum_{r=1}^{n} v_r(t+1) , \qquad (6)$$

где n — количество интервалов x_r ($r=\overline{1,n}$); τ_r — центроид r -го интервала по формуле (5); $v_r(t+1)$ — количество повторений НМ W_r в правой части ГНЛЗ, описывающей прогноз для (t+1) -го отсчета времени.

Средняя относительная ошибка прогнозирования вычисляется как

$$\sum_{t=1}^{m} \left| (f(t) - d(t)) / d(t) \right|$$

$$AFER = \frac{t=1}{m} \cdot 100\% , \qquad (7)$$

где f(t) и d(t) — предсказанное и реальное значения показателя для t -го отсчета времени; m — количество значений BP (количество отсчетов времени).

Для базовой модели С.М. Чена в работе были предложены следующие модификации:

- модели прогнозирования на основе ДНМТ1 с дополнительным неопределенным параметром с учетом и без учета повторов НМ в правых частях ГНЛЗ, в соотношениях (2) вместо степени принадлежности 0,5 используется число $\alpha \in [0,1]$, например $A_2 = \alpha/x_1 + 1/x_2 + \alpha/x_3 + 0/x_4 + ... + 0/x_n$;
- модели прогнозирования на основе непрерывных HMT1 (HHMT1) с учетом и без учета повторов HM в правых частях ГНЛЗ, в которых в представлении (1) вместо ДНМТ1 используются HHMT1 с гауссовскими $\Phi\Pi$:

 $u(x) = exp(-(x-b)^2/c^2),$ (8)

где координата максимума b ставится в соответствие середине интервала x_r ($r = \overline{1,n}$), а коэффициент концентрации c характеризует нечеткость HHMT1;

 модели прогнозирования на основе интервальных дискретных нечетких множеств второго типа (ИДНМТ2) и интервальных непрерывных нечетких множеств второго типа (ИННМТ2) без учета повторов НМ в правых частях ГНЛЗ, в которых в представлении (1) вместо НМТ1 используются ИДНМТ2 и ИННМТ2 с гауссовскими ФП.

В последнем случае ГНЛЗ вида (3) записывается как $FOU_j \rightarrow FOU_{l_1}$, FOU_{l_2} ,..., FOU_{l_g} , где FOU — «отпечаток неопределенности» («footprint of uncertainty»), а формула (4) приобретает вид:

$$u_{FOU}(x_r) = max(u_{FOU_{l_1}}(x_r), u_{FOU_{l_2}}(x_r), \dots, u_{FOU_{l_g}}(x_r)) \ (r = \overline{1,n}).$$

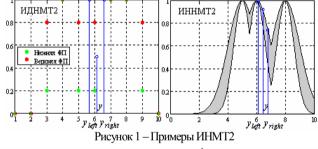
Центроид (дефаззифицированное значение) ИНМТ2 находится с помощью операции «понижения типа» («type-reduction») с использованием итерационного алгоритма Карника – Менделя²:

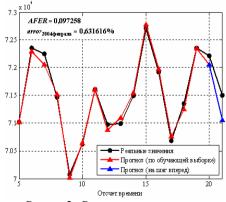
$$c_{\widetilde{A}} = \int \dots \int \int \dots \int 1/(\sum_{r=1}^n (w_r(t+1) \cdot \chi_r) / \sum_{r=1}^n w_r(t+1)) = [y_{left}, y_{right}],$$
 где $Z_r - \text{HMT1}$ с центром g_r и протяженностью $s_r (s_r \ge 0)$; $w_r - \text{HMT1}$ с центром h_r и протяженностью $\Delta_r (\Delta_r \ge 0)$; $r = \overline{1,n}$, и операции дефаззификации: $y(t+1) = (y_{left} + y_{right}) / 2$.

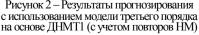
На рисунке 1 приведены примеры ИНМТ2 (при этом отмечены минимально и максимально возможные центроиды «вложенных» НМТ1 y_{left} и y_{right} , а также искомый центроид ИНМТ2 — y).

Для всех моделей прогнозирования должно выполняться требование об отсутствии ГНЛЗ с пустыми правыми частями для обеспечения возможности прогнозирования.

Самостоятельной задачей при прогнозировании с использованием моделей на основе НМТ1 и ИНМТ2 является определение оптимальных значений параметров моделей, обеспечивающих минимальное значение средней относительной ошибки прогнозирования *АFER*. Для решения задачи поиска оптимальных параметров указанных выше моделей прогнозирования были предложены модификации классического ГА. Использование ГА для поиска оптимальных параметров моделей прогнозирования позволяет улучшить результаты прогнозирования до 2-5 раз.







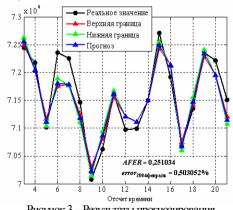


Рисунок 3 – Результаты прогнозирования с использованием модели первого порядка на основе ИДНМТ2 (без учета повторов НМ)

² Karnik N.N., Mendel J. M. Centroid of a type-2 fuzzy set // Information sciences, 2001. – Vol. 132. – P. 195-220.

	Прогне	оз, тыс. чел. / Относительная ошибка прогнозирования, %			
Реальное	Модель	Авторегрес-	Полино-	Модель	Модель
значение	Бурга	сионная	миальная	на основе	на основе
	Бурга	модель	модель	ДНМТ1	ИДНМТ2
71502	70900,609/	71197,104/	72366,353/	71050,382/	71146,042/
тыс. чел.	0,841083%	0,426416%	1,208851%	0,631616%	0,503052%

Кодирование хромосомы осуществляется следующим образом: для модели на основе ДНМТ1: $s=(D_1,D_2,n,k)$; для модели на основе ДНМТ1 с дополнительным неопределенным параметром: $s=(D_1,D_2,n,k,\alpha)$; для модели на основе ННМТ1: $s=(D_1,D_2,n,k,\sigma)$; для модели на основе ИДНМТ2: $s=(D_1,D_2,n,k,\alpha_{upper},\alpha_{lower})$; для модели на основе ИННМТ2: $s=(D_1,D_2,n,k,\sigma_{upper},\sigma_{lower})$, где D_1 , D_2 — числа корректировки границ универсума; n — количество интервалов разбиения универсума; k — порядок модели; α — степень принадлежности в модели на основе ДНМТ1 (α \in [0,1]); $\alpha_{upper},\alpha_{lower}$ — степени принадлежности для «верхней» и «нижней» ФП ДНМТ1, формирующих ИДНМТ2 ($\alpha_{upper}>\alpha_{lower}$); $\sigma_{upper},\sigma_{lower}$ — коэффициенты концентрации для «верхней» и «нижней» гауссовских ФП ННМТ1, формирующих ИННМТ2 ($\sigma_{upper}>\sigma_{lower}$).

При формировании начальной популяции и выполнении операций скрещивания и мутации для моделей прогнозирования на основе ИДНМТ2 (ИННМТ2) должна выполняться проверка условия: $\alpha_{upper} > \alpha_{lower}$ ($\sigma_{upper} > \sigma_{lower}$).

Хромосома, для которой не выполняется требование об отсутствии ГНЛЗ с пустыми правыми частями, считается «нежизнеспособной» и ее функция соответствия полагается равной числу 100. В связи с этим в качестве функции соответствия в модели прогнозирования на основе НМТ1 (ДНМТ1 и ННМТ1) используется функция вида:

$$FF_{\text{HMT1}}(s) = \begin{cases} AFER, & \text{если определены все правые части} \\ & \text{в группах логических зависимостей} \\ & \text{первого порядка и } k$$
-го порядка 100, $& \text{если не определена хотя бы одна правая} \\ & \text{часть в группах логических зависимостей} \\ & \text{первого порядка} \quad \text{и } k$ -го порядка,
$$\end{cases}$$

в модели прогнозирования на основе ИДНМТ2 – функция вида:

$$FF_{\text{ИНМТ2}}(s) = \begin{cases} AFER & \text{при } AFER < FF_{\alpha_{upper}}, & AFER < FF_{\alpha_{lower}} \\ 100 & \text{иначе,} \end{cases}$$
 (10)

где $FF_{\Lambda}(s) = FF_{\mathrm{HMT1}}(s)$ при $\Lambda = \alpha_{lower}, \alpha_{upper}$.

Для модели прогнозирования на основе ИННМТ2 выполняется замена α_{lower} , α_{upper} на σ_{lower} , σ_{upper} в формуле (10).

На рисунках 2 и 3 приведены графические зависимости для реальных и прогнозируемых значений показателя «экономически активное население», при этом прогнозирование выполнялось на основе поквартальных данных («1999 февраль» – «2003 ноябрь») для отсчета времени «2004 февраль».

В таблице 1 приведены результаты прогнозирования для отсчета времени «2004 февраль» с использованием классических моделей (Бурга, авторегрессионной и полиномиальной) и моделей на основе НВР. Относительная ошибка прогнозирования вычислялась как $error = |y(m+1) - d(m+1)|/d(m+1) \cdot 100\%$, где y(m+1) и d(m+1) — предсказанное и реальное значения показателя для (m+1) -го отсчета времени соответственно. Анализ результатов показывает, что предлагаемые модели позволяют получать результаты прогнозирования, близкие к результатам классических моделей, а в ряде случаев и лучшие, обеспечивая при этом простоту реализации и наглядность представления процесса про-

гнозирования, не требуя от специалиста-аналитика выполнения расчета группы критериев качества модели с последующим их анализом для выбора оптимальных параметров модели, как в случае работы с классическими моделями. Результаты прогнозирования, полученные для показателей, характеризующих уровень занятости населения в экономике страны, внешнеторговую деятельность региона и др., свидетельствуют о пригодности разработанных моделей прогнозирования для выполнения краткосрочных прогнозов. Разработанные модели прогнозирования обеспечивают расширение информационноаналитического пространства при решении задач краткосрочного прогнозирования за счет представления значений ВР с помощью НМТ1 и ИНМТ2.

В главе 3 «Оценка, классификация и упорядочение объектов с использованием нечеткого метода Дельфы, мультимножеств, систем нечеткого вывода и генетических алгоритмов» рассматриваются методы оценки, классификации и упорядочения объектов при наличии несовпадающих, в том числе противоречивых, исходных данных (экспертных оценках).

В п. 3.1 предлагается метод оценивания объектов по критериям при наличии несовпадающих экспертных оценок. Рассматривается оценивание объектов с помощью треугольных нечетких чисел, представляющих собой тройку чисел: минимальное, наиболее правдоподобное и максимальное значения: $A_r = (a_1^r, a_M^r, a_2^r)$ (где $r = \overline{1,n}$, n – количество экспертов), на основе которых в ходе нескольких туров согласования с использованием нечеткого метода Дельфы определяются «частично» согласованные треугольные нечеткие числа экспертов и среднее треугольное число $A_{average} = (m_1, m_M, m_2)$, в результате дефаззификации которого по методу центра тяжести определяется оценка — центроид НМТ1 (рисунок 4). Если объекты имеют одинаковые оценки на основе центроидов НМТ1, предлагается использовать центроиды ИНМТ2, вычисляемые с использованием итерационного алгоритма Карника – Менделя и позволяющие в ряде случаев дифференцировать объекты по критериям оценивания. Предлагаемый метод оценивания был использован при решении залачи оценки недвижимости как залачи оценки потенциального качества изменяемых многомерных объектов при малых объёмах априорной информашии.



В пп. 3.2 и 3.3 рассматриваются вопросы разработки систем поддержки принятия решений с использованием систем нечеткого вывода на основе НМТ1 и ИНМТ2; исследуются наиболее часто используемые алгоритмы нечеткого вывода – алгоритмы Мамдани и Сугено; предлагаются ГА для выполнения «тонкой» настройки параметров систем нечеткого вывода на основе НМТ1 и ИНМТ2 при использовании алгоритмов Мамдани и Сугено с применением обучающих выборок, решающие проблему попадания в локальный минимум, возникающую при использовании классических методов оптимизации (типа градиентного спуска) целевой функции.

В качестве ФП термов лингвистических переменных (ЛП) предлагается использовать гауссовские $\Phi\Pi$ (формула (8)), которые зависят только от двух параметров: b — координаты центра максимума, с - коэффициента концентрации ФП, что позволяет минимизировать длину хромосомы, а следовательно, время поиска оптимального решения. При реализации ГА может выполняться настройка параметров ФП входных и выходных ЛП и весовых коэффициентов правил нечеткого вывода.

Пусть y=Z(X,C,B,W) — модель на основе HMT1, соответствующая системе нечеткого вывода, где $X=(x_1,...,x_q)$ — вектор входных ЛП; $C=(c_{1,1},...,c_{1,e_1},...,c_{q,1},...,c_{q,e_q})$ и $B=(b_{1,1},...,b_{1,e_1},...,b_{q,1},...,b_{q,e_q})$ — векторы параметров гауссовских ФП для термов входных ЛП: координаты центров максимумов и коэффициенты концентрации соответственно (при этом параметры термов выходной ЛП не настраиваются); $W=(w_1,...,w_N)$ — вектор весов правил из нечеткой базы знаний; N — количество правил в нечеткой базе знаний; e_i — количество термов ЛП x_i . Пусть обучающая выборка задана в виде M пар экспериментальных данных: (X^v,y^v) , $v=\overline{1,M}$, где $X^v=(x_1^v,...,x_q^v)$, y^v — вектор значений входных ЛП и соответствующее значение выходной ЛП y для v -й пары «входывыход». В соответствии с методом наименыших квадратов задача оптимальной настройки модели на основе HMT1 формулируется в следующем виде: найти вектор (C,B,W), удовлетворяющий ограничениям: $c_{i,r} \in [\underline{c}_{i,r}, \overline{c}_{i,r}]$, $b_{i,r} \in [\underline{b}_{i,r}, \overline{b}_{i,r}]$, $w_j \in [\underline{w}_j, \overline{w}_j]$ $(i=\overline{1,q}, r=\overline{1,e_i}, j=\overline{1,N})$ и обеспечивающий минимум целевой функции:

$$F(C, B, W) = \sum_{v=1}^{M} [Z(X^{v}, C, B, W) - y^{v}]^{2}.$$
 (11)

Для реализации ΓA хромосома определяется в виде:

$$s = (C, B, W) = (c_{1,1}, b_{1,1}, ..., c_{1,e_1}, b_{1,e_1}, c_{q,1}, b_{q,1}, ..., c_{q,e_q}, b_{q,e_q}, w_1, ..., w_N),$$
 где N – количество правил вывода; e_i – количество термов входной ЛП x_i .

В качестве функции соответствия для ГА выбирается функция вида (11). Особенностью реализации ГА является необходимость контроля за упорядоченностью термов входных ЛП: $c_{i,r} < c_{i,r+1}$ ($i=\overline{1,q}$, $r=\overline{1,e_i-1}$) при создании начальной популяции хромосом и выполнении операций скрещивания и мутации (если данное требование не выполняется, реализуется новая попытка генерации хромосомы или выполнения генетических операций). Одноточечное скрещивание выполняется для каждой входной ЛП x_i ($i=\overline{1,q}$) и для весовых коэффициентов правил w_k ($k=\overline{1,N}$). Мутации может подвергаться не более 10-20 % генов в хромосоме.

Аналогичным образом реализуется настройка параметров систем нечеткого вывода на основе ИНМТ2. При этом дополнительно для коэффициентов концентрации «верхней» и «нижней» ФП термов входных ЛП выполняется проверка условия: $b_{i,r}^{upper} > b_{i,r}^{lower}$ ($i = \overline{1,q}$, $r = \overline{1,e_i}$). Использование ИНМТ2 позволяет управлять неопределенностью при выборе параметров модели (за счет вовлечения в решение задачи большего количества

информации) и повысить точность аппроксимации на обучающей выборке.

Предлагается использовать системы нечеткого вывода как самостоятельные системы поддержки принятия решений и как вспомогательный инструментарий при решении других прикладных задач. Предлагаемые подходы к разработке систем нечеткого вывода были использованы при реализации иерархической системы нечеткого вывода (с одновременным использованием алгоритмов Мамдани и Сугено, обеспечивающим адекватность и обоснованность принимаемых решений) для решения задачи диагностики состояния ГИК и принятия решений в штатных и аварийных ситуациях; задачи классификации способов несанкционированного отбора электроэнергии с выполнением «тонкой» настройки параметров системы нечеткого вывода на основе НМТ1 и ИНМТ2.

На рисунке 5 показаны примеры поверхностей систем нечеткого вывода для классификации способов несанкционированного отбора электроэнергии на основе «среднего» (статистического) и «текущего» потребления электроэнергии абонентов электросети на пять классов («Норма»; «Магнит»; «Частичный обход прибора учета электроэнергии»; «Обход прибора учета электроэнергии»; «Скругка электросчетчика»), демонстрирующие эффективность выполнения «тонкой» настройки параметров системы нечеткого вывода.

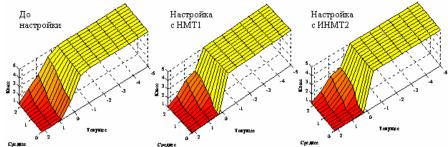


Рисунок 5 – Поверхности нечеткого вывода

Таблица 2

Оцениваемые характеристики	До настройки	Настройка с НМТ1	Настройка с ИНМТ2
Значение функции соответствия	48,1324	4,2776	2,9615
Ошибка классификации на обучающей выборке (при M=50)	39	2	0

В таблице 2 приведены оценки качества аппроксимации на обучающей выборке для систем нечеткого вывода до и после выполнения «тонкой» настройки.

В пп. 3.4 и 3.5 рассматриваются вопросы многокритериальной классификации и упорядочения объектов при групповом экспертном оценивании, характеризующемся наличием несовпадающих, в том числе противоречивых, экспертных оценок. Предлагается при классификации объектов применять прямой метод оценивания объектов по критериям, а при упорядочении объектов внутри класса — выполнять оценивание с использованием метода парных сравнений по шкале Саати.

Пусть $X = \{x_1, x_2, ..., x_n\}$ — множество объектов; $P = \{P_1, P_2, ..., P_q\}$ — множество количественных и качественных критериев, m — количество экспертов. В качестве математической модели для представления объектов, оцениваемых группой экспертов по ряду критериев, выбрано мультимножество, использование которого позволяет учесть все комбинации значений оценок по количественным и качественным критериям, а также количество значений оценок по каждому из этих критериев (не прибегая к дополнительным преобразованиям типа усреднения, смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных). В этом случае возможен учет всех, в том числе и противоречивых, оценок по критериям для описания объектов.

Каждый эксперт выполняет предварительную индивидуальную сортировку объектов x_i на несколько классов X_v ($v=\overline{1,f}$) путем прямой классификации. Принадлежность объекта x_i к классу выражается индивидуальным правилом сортировки $W=\{w_v\}$, которое является дополнительным качественным критерием. Пусть по каждому l-му критерию экспертами дано z_l различных оценок $p_l^{y_l}$ ($y_l=\overline{1,z_l}$); количество экспертов, давших объекту x_i оценку $p_l^{y_l}$, равно $k_{x_i}(p_l^{y_l})$; количество экспертов, давших оценку w_v ,

равно
$$k_{x_i}(w_v)$$
 ($\sum_{y_i=1}^{z_l} k_{x_i}(p_l^{y_l}) = m$, $\sum_{v=1}^f k_{x_i}(w_v) = m$, $i=\overline{1,n}$, $l=\overline{1,q}$).

В п. 3.4 при выполнении предварительных индивидуальных сортировок объектов экспертами предлагается использовать индивидуальные системы нечеткого вывода, параметры которых настроены с помощью ГА на основе экспертных обучающих выборок. Использование систем нечеткого вывода позволяет не только формализовать знания экспертов в конкретной прикладной области, но и обеспечить формирование экспертных классификационных решений в том случае, когда эксперт затрудняется на основе вы-

ставленных им же оценок по критериям выполнить предварительную сортировку (классификацию) объекта. Исходя из предположений, что: системы нечеткого вывода являются универсальными аппроксиматорами; переход от принадлежности к классу к непринадлежности является плавным, а не скачкообразным, удается решить проблему предварительной экспертной классификации объектов. Системы нечеткого вывода, реализующие классификацию, ИП были использованы при выполнении предварительных индивидуальных сортировок ИП при их групповом экспертном оценивании.

Объект x_i существует в m экземплярах, отличающихся наборами значений оценок по множеству критериев P, и имеется m несовпадающих индивидуальных сортировок множества объектов X. В результате формируется расширенное множество критериев: $G = \{P_1, P_2, ..., P_q, W\}$, при этом в описании каждого экземпляра объекта x_i присутствует только одно значение оценки по критерию из каждой группы расширенного множества критериев. Каждому объекту x_i в соответствие ставится мультимножество s_i вида:

$$s_{i} = \{k_{x_{i}}(p_{1}^{1}) \bullet p_{1}^{1}, ..., k_{x_{i}}(p_{1}^{z_{1}}) \bullet p_{1}^{z_{1}}, ..., k_{x_{i}}(p_{q}^{1}) \bullet p_{q}^{1}, ..., k_{X_{v}}(p_{q}^{z_{q}}) \bullet p_{q}^{z_{q}}, k_{x_{i}}(w_{1}) \bullet w_{1}, ..., k_{x_{i}}(w_{f}) \bullet w_{f}\},$$

$$(13)$$

где $k_{x_i}(p_l^{y_l})$ и $k_{x_i}(w_v)$ — количество экспертов, давших объекту x_i оценки $p_l^{y_l}$ и w_v ; «•»—знак кратности вхождения оценки $p_l^{y_l}$ в мультимножество; $y_l=\overline{1,z_l}$; v=1,f.

Для мультимножеств вида (13) формируются обобщенные решающие правила классификации, составленные из небольшого числа значений по критериям, относящих объекты к заданным классам наилучшим образом в смысле близости к предварительным индивидуальным сортировкам. Представление объекта x_i в таком виде является способом выражения индивидуальных правил сортировки: «ЕСЛИ <условия», ТО <решение»». Терм <условия» определяет комбинации оценок критериев $p_l^{y_i}$, описывающих объект x_i , а терм <решение» – принадлежность объекта x_i классу X_v .

В терм <решение> входят совокупности индивидуальных заключений экспертов по предварительной сортировке объекта x_i и правило (например, правило простого большинства голосов или квалифицированного большинства голосов), относящее объект x_i к классу X_{v} . При формировании каждого класса реализуется сложение входящих в него мультимножеств с учетом оценок по критериям всех объектов, входящих в класс. Объекты, попавшие в разложение на классы по результатам предварительных индивидуальных сортировок, образуют наилучшую из всех возможных декомпозиций множества объектов.

При определении расстояния между мультимножествами A и B рассматривается метрическое пространство мультимножеств (S,d) с метрикой Хемминга:

$$d = m(A\Delta B) = \sum_{l=1}^{q} \chi_l \cdot \sum_{y_l=1}^{z_l} \left| k_A(p_l^{y_l}) - k_B(p_l^{y_l}) \right|,$$
 (14)

где $\chi_l > 0$ $(l = \overline{1,q})$ – коэффициенты относительной важности критериев.

Поиск обобщенных решающих правил многокритериальной классификации объектов сводится к q оптимизационным задачам по каждой группе критериев P_l : необходимо найти такие f мультимножеств, которые будут располагаться на максимально возможном расстоянии в пространстве мультимножеств и принадлежать разным классам. Например, при f=2 вычисляются значения $p^*_l(v_1,v_2)$ оценок по критериям, определяющие границы между классами v_1 и v_2 ($v_1,v_2=\overline{1,f}$, $v_1\neq v_2$). Для построения обобщенных решающих правил классификации используются граничные значения критериев $p^*_l(v_1,v_2)$, которым соответствуют максимальные значения расстояний $d_l(v_1,v_2)$.

Оценка точности аппроксимации по l-й группе критериев вычисляется как $\rho_l = d_l(v_1, v_2)/d(v_1, v_2)$, где $d(v_1, v_2)$ — расстояние между классами v_1 и v_2 , определенное с помощью предварительных индивидуальных сортировок объектов. Величина ρ_l характеризует относительную важность l-й группы критериев P_l . Обобщающие решающие правила классификации включают граничные значения критериев $p^*_l(v_1, v_2)$, имеющие показатель точности, превышающий желаемый пороговый уровень ρ_0 .

При выполнении многокритериального упорядочения объектов x_i ($i = \overline{1,n}$) формируются мультимножества вида:

$$s_{i} = \{k_{x_{i}}(p_{1}^{1}) \bullet p_{1}^{1}, ..., k_{x_{i}}(p_{1}^{z_{1}}) \bullet p_{1}^{z_{1}}, ..., k_{x_{i}}(p_{q}^{1}) \bullet p_{q}^{1}, ..., k_{X_{u}}(p_{q}^{z_{q}}) \bullet p_{q}^{z_{q}}\} .$$

$$(15)$$

Упорядочение объектов производится по близости к гипотетическим «идеальному» объекту: $s_{max} = \{m \bullet p_1^1, 0, ..., 0, m \bullet p_2^1, 0, ..., 0, m \bullet p_q^1, 0, ..., 0\}$ или «антиидеальному» объекту: $s_{min} = \{0, ..., 0, m \bullet p_1^{\bar{z}_1}, 0, ..., 0, m \bullet p_2^{\bar{z}_2}, 0, ..., 0, m \bullet p_q^{\bar{z}_q}\}$. Задача многокритериального упорядочения объектов по близости к «идеальному» объекту сводится к сравнению взвешенных сумм первых оценок объектов по критериям P_l ($l = \overline{1,q}$): $H_{s_h}^1 = \sum_{l=1}^q \chi_l \cdot k_{x_h}(p_l^1)$. Лучшим будет объект x_h , для которого сумма $H_{s_h}^1$ будет наибольшей. Для упорядочения объектов внутри группы эквивалентных объектов x_h ($t = \overline{1,r}$), имеющих равные суммы $H_{s_h}^1$, вычисляются взвешенные суммы вторых оценок объектов по критериям P_l : $H_{s_{h_l}}^2 = \sum_{l=1}^q \chi_l \cdot k_{x_{h_l}}(p_l^2)$. Процесс вычисления взвешенных сумм повторяется до полного упорядочения всех объектов.

Объекты, для которых выполняется многокритериальное упорядочение, оцениваются с использованием метода парных сравнений по шкале Саати. Аналогичным образом выполняется оценивание относительной степени важности критериев. Интегральная оценка объектов вычисляется в соответствии со схемой Беллмана — Заде. Пусть $u_{P_i}(x_i)$ — число в диапазоне [0,1], характеризирующее уровень оценки объекта $x_i \in X$ по критерию $P_l \in P$: чем больше число $u_{P_l}(x_i)$, тем выше оценка объекта x_i по критерию P_l , $i=\overline{1,n}$, $l=\overline{1,q}$. Критерий P_l представляется в виде НМ \widetilde{P}_l на множестве объектов X:

$$\widetilde{P}_{l} = \left\{ \frac{u_{P_{l}}(x_{1})}{x_{1}}, \quad \frac{u_{P_{l}}(x_{2})}{x_{2}}, \quad ..., \quad \frac{u_{P_{l}}(x_{n})}{x_{n}} \right\},$$
 где $u_{P_{l}}(x_{i})$ определяется с использованием мат-

рицы парных сравнений $A_{n\times n}$: $u(x_i)=1/\sum_{k=1}^n a_{ki}$. В общем случае нечеткое решение \widetilde{D} находится как пересечение частных критериев:

$$\widetilde{D} = \widetilde{P}_1 \cap \widetilde{P}_2 \cap \dots \cap \widetilde{P}_q = \left\{ \frac{\min_{l=1,q} (u_{P_l}(x_1))^{\chi_l}}{x_1}, \frac{\min_{l=1,q} (u_{P_l}(x_2))^{\chi_l}}{x_2}, \dots, \frac{\min_{l=1,q} (u_{P_l}(x_n))^{\chi_l}}{x_n} \right\},$$
(16)

где χ_l — коэффициент относительной важности критерия P_l ($\sum_{l=1}^q \chi_l = 1$).

Лучшим объектом является тот, для которого степень принадлежности является наибольшей, на втором месте будет объект, для которого значение ФП занимает второе место в списке упорядочения по убыванию и т.д. дем называть min -решениями. Эксперты могут дать различные, в том числе и противоречащие друг другу, решения задачи упорядочения объектов. Пусть каждый r -й эксперт в соответствии с формулой (16) упорядочил объекты x_i из множества X по критериям из множества P и получил нечеткое решение \widetilde{D}_r . Для повышения адекватности результатов упорядочения объектов предлагается выполнение нескольких этапов согласования экспертных оценок с использованием нечеткого метода Дельфы для оценок, полученных методом парных сравнений. Для каждого r-го эксперта строится индивидуальная «обобщенная» матрица парных сравнений B_r по всем критериям, элементы которой вычисляются как $b_{ij}=d_r(x_i)/d_r(x_j)$, где $d_r(x_i)=\min_{l=1,q}(u_{r,P_l}(x_i))^{\chi_l}$ $(i,j=\overline{1,n}\;,\;r=\overline{1,m}\;)$. На основе матриц B_r ($r = \overline{1, m}$) вычисляется матрица C средневзвешенных согласованных оценок с элементами: $c_{ij} = \left(\prod_{i=1}^m (b_{ij,r})^{\beta_r}\right)^{1/m}$, где β_r — коэффициенты относительной важности экспертов. Если на некотором α -уровне ($\alpha \in [0,1]$), то есть при $c_{ii} > M \cdot \alpha$ $(i,j=\overline{1,n}\,)$: $(b_{ij,r})^{eta_r} < c_{ij}\,$, то $\,r\,$ -й эксперт должен обосновать оценки, выставленные им по отдельным критериям P_l ($l=\overline{1,q}$) для объектов x_i и x_i , и, возможно, и предложенные им коэффициенты относительной важности критериев, так как он оценил уровень преимущества объекта x_i над x_i ниже, чем все остальные эксперты. Если все эксперты согласны выполнить новое оценивание объекта (новый тур оценивания), то осуществляется вычисление новых значений матриц парных сравнений A по различным критериям для каждого эксперта, на основе которых находятся новые нечеткие решения \widetilde{D}_r $(r=\overline{1,m})$. Правило завершения процедуры согласования имеет вид $(b_{ii},r)^{\beta_r}>s_{ii}-\gamma$, где $i, j = \overline{1, n}, r = \overline{1, m}, \gamma$ — балл, определяющий допустимые отклонения оценки эксперта от усредненного балла. Итерационный процесс согласования экспертных оценок повторяется до тех пор, пока оценки экспертов не сойдутся к достаточно узкому интервалу либо пока хотя бы один из экспертов не откажется от выполнения процедуры согласования оценок. В качестве окончательного нечеткого решения (min -решения) задачи многокритериального упорядочения объектов группой экспертов принимается решение, являющееся пересечением частных решений экспертов:

Интегральные нечеткие решения, определяемые в соответствии с формулой (16), бу-

$$\widetilde{D} = \widetilde{D}_1 \cap \widetilde{D}_2 \cap ... \cap \widetilde{D}_r = \left\{ \frac{\min_{r=1,m} d_r(x_1)}{x_1}, \frac{\min_{r=1,m} d_r(x_2)}{x_2}, \dots, \frac{\min_{r=1,m} d_r(x_n)}{x_n} \right\}.$$

$$(17)$$

Наряду с нечетким решением по упорядочению объектов в соответствии с формулой (17) следует рассмотреть альтернативный вариант упорядочения на основе представления объектов в виде мультимножеств. Если эксперты не согласны выполнить новое оценивание объектов, а желаемый уровень согласованности оценок экспертов не достигнут, то представление объектов с помощью мультимножеств позволяет упорядочить оптимальным (наилучшим) образом объекты по близости к «идеальному» («антиидеальному») объекту даже при наличии полностью противоречивых оценок экспертов. Таким образом, в конечном итоге будут получены 2 варианта решения задачи многокритериального упорядочения объектов: классический вариант – в соответствии с формулой (17); альтернативный вариант – с использованием представления объектов в виде мультимножеств по формуле (15). Классический вариант упорядочения реализует стратегию «лучше недооценить, чем переоценить» (так как из-за применяемой операции взятия минимума будут «уграчены» все «лучшие» экспертные оценки), использование которой

при решении некоторых практических задач, например при оценке финансовой состоятельности ИП, нельзя недооценивать, а альтернативный — учитывает все, в том числе и противоречивые, оценки по критериям для принятия решения об упорядочении объектов.

Глава 4 «Методы и алгоритмы кластеризации на основе нечетких множеств первого типа и генетических алгоритмов» посвящена исследованию алгоритмов кластеризации на основе НМТ1 и разработке эффективных методов кластеризации, реализующих комбинирование классических алгоритмов кластеризации и ГА, обеспечивающих получение адекватных результатов кластеризации множества объектов, содержащего кластеры подобной плотности и подобного объема, с минимально возможными временными затратами.

Пусть n — количество объектов кластеризации, q — количество критериев оценивания, c — количество кластеров. Базовый алгоритм четких c -средних обеспечивает получение адекватных результатов кластеризации, если кластеры компактны и хорошо отделимы. В противном случае следует использовать алгоритмы кластеризации на основе HMT1, реализующие учет того или иного вида неопределенности: кластеризации объектов может выполняться c учетом свойств кластерной относительности и кластерной типичности. Рассмотрены классический алгоритм нечеткой кластеризации — алгоритм нечетких c -средних (FCM-алгоритм на основе HMT1; J.C. Bezdek, J.C. Dunn), реализующий учет свойства кластерной относительности, и его модификации: алгоритм возможностных c -средних (PCM-алгоритм на основе HMT1; J. Keller, R. Krishnapuram) и алгоритм возможностных нечетких c -средних (PFCM-алгоритм на основе HMT1; R.N. Dave, R. Krishnapuram), обеспечивающие учет свойства кластерной типичности, а также робастный алгоритм нечетких c -средних (RFCM-алгоритм на основе HMT1; Y. Ohashi, R.N. Dave), локализующий объекты-шумы в отдельный виртуальный кластер шума.

FCM-алгоритм на основе HMT1 реализует минимизацию целевой функции:

$$J(U,V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (u_j(x_i))^m \cdot d_{ji}^2 \left(d_{ji}^2 = \sum_{l=1}^{q} (x_i^l - v_j^l)^2 \right)$$
 (18)

при

$$\sum_{j=1}^{c} u_{j}(x_{i}) = 1 \ (\forall i = \overline{1, n}), \tag{19}$$

где $U=[u_j(x_i)]$ — нечеткое c -разбиение множества объектов $\{x_i\}$ на основе $\Phi\Pi$ $u_j(x_i)$; $V=(v_1,...,v_c)$ — центры кластеров; d_{ji} — евклидово расстояние между центром кластера v_j и объектом x_i : $d_{ji}=\mid x_i-v_j\mid ; m$ — фаззификатор ($m\in R$, m>1); c — количество кластеров; n — количество объектов.

FCM-алгоритм предполагает выполнение следующих шагов.

- 1. Инициализация начального нечеткого разбиения $U = [u_i(x_i)]$.
- 2. Вычисление центров кластеров:

$$v_{j} = \sum_{i=1}^{n} (u_{j}(x_{i})^{m} \cdot x_{i}) / \sum_{i=1}^{n} u_{j}(x_{i})^{m} .$$
(20)

3. Вычисление новых ФП, определяемых через относительные расстояния между объектами и центрами кластеров:

$$u_{j}(x_{i}) = 1/\sum_{k=1}^{c} \left(d_{ji}/d_{ki}\right)^{2/(m-1)}.$$
(21)

4. Шаги 2^{k-1} 3 повторяются до тех пор, пока не будет выполнено заданное число итераций s или не будет достигнута заданная точность: $|J(U,V)-J'(U,V)| \le \varepsilon$, где J(U,V), J'(U,V) — значения целевой функции на двух последовательных итерациях.

В п. 4.2 выполнен анализ, показавший целесообразность использования в качестве показателя качества кластеризации с использованием FCM-алгоритма на основе HMT1 индекса Ce – Бени (X.L. Xei, G.A. Beni), характеризующего компактность и отделимость кластеров гиперсферической формы:

$$XB = \frac{\sum_{j=1}^{c} \sum_{i=1}^{n} (u_{j}(x_{i}))^{2} \cdot ||x_{i} - v_{j}||^{2}}{n \cdot \min_{t \neq j} ||v_{t} - v_{j}||^{2}},$$
(22)

а для кластеров гиперэллиптической формы – нечеткого общего гиперобъема:

$$FH = \sum_{j=1}^{c} (det(R_j))^{1/2} \left(R_j = \sum_{i=1}^{n} ((u_j(x_i)^m \cdot (x_i - v_j) \cdot (x_i - v_j)^T) / \sum_{i=1}^{n} u_j(x_i)^m \right), \tag{23}$$

где c – количество кластеров; n – количество объектов; $u_i(x_i)$ – $\Phi\Pi$ объекта x_i кластеру X_i ; v_i – вектор координат центра j -го кластера; x_i – вектор координат i -го объекта; $j = \overline{1,c}$; $t = \overline{1,c}$; R_j — нечеткая ковариационная матрица j -го кластера; $det(R_i)$ — определитель матрицы R_i .

Оба показателя качества кластеризации должны быть минимизированы.

При хороших результатах нечеткой кластеризации XB < 1.

В случае возможностной интерпретации неопределенности может быть использован РСМ-алгоритм на основе НМТ1, позволяющий улучшить результаты кластеризации за счет ослабления свойства кластерной относительности и учета свойства кластерной типичности, если множество объектов содержит атипичные объекты.

РСМ-алгоритм основан на минимизации целевой функции вида

$$J(W,V) = \sum_{i=|i|=1}^{c} \sum_{j=|i|=1}^{n} (w_j(x_i))^m \cdot d_{ji}^2 + \sum_{i=1}^{c} \eta_j^2 \cdot \sum_{i=1}^{n} (1 - w_j(x_i))^m , \qquad (24)$$

 $J(W,V) = \sum_{j=li=1}^{c} \sum_{i=li=1}^{n} (w_j(x_i))^m \cdot d_{ji}^2 + \sum_{j=1}^{c} \eta_j^2 \cdot \sum_{i=1}^{n} (1-w_j(x_i))^m$, (24) где $W = [w_j(x_i)]$ — возможностное c -разбиение множества объектов $\{x_i\}$ на основе функций типичности (ФТ) $w_j(x_i)$; $V = (v_1,...,v_c)$ — центры кластеров; d_{ji} — расстояние между центром кластера v_j и объектом x_i ; m — фаззификатор; η_i ($j = \overline{1,c}$) — «ширина зоны», определяющая расстояние, на котором значение ФТ объекта / -му кластеру равно 0,5; с - количество кластеров; п - количество объектов кластеризации.

При кластеризации с использованием РСМ-алгоритма на основе НМТ1 отбрасывается ограничение вида (19), которое обеспечивает выполнение требования о том, что для любого объекта x_i ($i = \overline{1, n}$) сумма степеней принадлежности всем кластерам должна равняться 1.

Возможностные ФТ, определяемые через абсолютное расстояние между центром кластера v_i и объектом x_i , могут быть вычислены как

$$w_j(x_i) = 1/(1 + (d_{ji}/\eta_j)^{2/(m-1)})$$
. При этом (25)

$$\sum_{i=1}^{n} w_{j}(x_{i}) = 1 \ (j = \overline{1, c}).$$
 (26)

 $\dot{\bar{\text{Koop}}}$ динаты центров кластеров находятся в соответствии с формулой (20).

Для уменьшения влияния атипичных объектов на результаты кластеризации может использоваться РГСМ-алгоритм на основе НМТ1, минимизирующий целевую функцию:

$$J(U,W,V) = \sum_{i=1}^{c} \sum_{i=1}^{n} (u_j(x_i)^m + (w_j(x_i))^{m'}) \cdot d_{ji}^2,$$
(27)

где m и m' – фаззификаторы ($m, m' \in R$, m, m' > 1), а $\Phi\Pi$ $u_i(x_i)$ и ΦT $w_i(x_i)$ рассчитываются по формулам (21) и (25) соответственно.

При этом координаты центров кластеров определяются как

$$v_j^l = \sum_{i=1}^n ((u_j(x_i))^m + (w_j(x_i))^{m'}) \cdot x_i^l / \sum_{i=1}^n ((u_j(x_i))^m + (w_j(x_i))^{m'}).$$
 (28)

RFCM-алгоритм на основе HMT1, обеспечивающий локализацию объектов-шумов в виртуальный кластер шума, основан на минимизации целевой функции вида

$$J(U,V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (u_j(x_i))^m \cdot d_{ji}^2 + \sum_{i=1}^{n} \delta^2 \cdot (1 - \sum_{j=1}^{c} u_j(x_i))^m , \qquad (29)$$

где δ — весовой коэффициент, характеризующий расстояние между множеством объектов и кластером шума.

Для каждого j -го кластера ($j = \overline{1,c}$) ФП находятся по формуле

$$u_{j}(x_{i}) = 1/\left(\sum_{r=1}^{c} (d_{ji}/d_{ri})^{-2/(m-1)} + \sum_{r=1}^{c} (d_{ji}/\delta)^{-2/(m-1)}\right)$$
, а для кластера шума ФП опре-

деляется как $u_*(x_i) = 1 - \sum_{i=1}^{c} u_i(x_i)$.

Инициализация начального разбиения в модификациях FCM-алгоритма на основе HMT1 выполняется с помощью нескольких итераций FCM-алгоритма.

«Ширина зоны» приблизительно может быть оценена как:

$$\eta_{j} = \left(\frac{K}{n_{j}} \cdot \sum_{i = 1}^{n} (u(x_{i}))^{m} \cdot d_{ij}^{2}\right)^{1/2},\tag{30}$$

а весовой коэффициент δ — как

$$\delta = (\lambda \cdot \sum_{j=1}^{c} \sum_{i=1}^{n} d_{ji}^{2} / (n \cdot c))^{1/2},$$
(31)

где $u_j(x_i)$ — $\Phi\Pi$, определяющая степень принадлежности объекта x_i кластеру X_j по формуле (21) для FCM-алгоритма; d_{ji} — расстояние между центром кластера v_j и объектом x_i ; n_j — количество объектов, отнесенных к j -му кластеру; c — количество кластеров; n — количество объектов; K — действительное число (обычно K = 1); λ — некоторый множитель, используемый для получения значения весового коэффициента δ .

Результаты кластеризации существенным образом зависят от правильности выбора значений параметров η_j и δ .

Для получения адекватных результатов кластеризации с использованием алгоритма кластеризации на основе HMT1, определяющего локально-оптимальное разбиение, требуется многократное выполнение этого алгоритма при заданном количестве кластеров c для различных исходных разбиений $R(X) = \{X_j \mid X_j \subseteq X\}$ с целью определения разбиения, которому соответствует минимальное значение целевой функции (показателя качества кластеризации) для принятия окончательного решения об искомом разбиении.

В п. 4.4 для получения адекватных результатов кластеризации с использованием FCM-алгоритма на основе НМТ1 при минимальных временных затратах предложен комбинированный метод нечеткой кластеризации (КМНК), реализующий поочередное выполнение FCM-алгоритма и ГА. При этом в качестве функции соответствия предлагается использовать показатель качества кластеризации по формуле (18) или (19).

Для заданного количества кластеров c хромосома может быть закодирована координатами центров всех кластеров или степенями принадлежности (числами из интервала $[0,\ 1]$) объектов центрам кластерам. При кодировании хромосомы центрами кластеров длина хромосомы равна $c\cdot q$, где c — числа кластеров, q — количество критериев: первые q координат соответствуют центру первого кластера, вторые q координат — центру второго кластера и т.п. При кодировании хромосомы степенями принадлежности объектов центрам нечетких кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$, где c — количество кластеров длина хромосомы равна $c\cdot n$ — количество кластеров длина хромосом дли хромосом дли хромосом

стеров, n — количество объектов: первые c элементов хромосомы соответствуют степеням принадлежности первого объекта центрам кластеров, вторые c элементов хромосомы — степеням принадлежности второго объекта центрам кластеров и т.п.

При формировании начальной популяции проверяется требование о разбиении множества объектов на заданное количество кластеров *с* (фактическое, а не формальное). Если данное требование не выполняется, то хромосома признается «нежизнеспособной» и отбрасывается. Таким образом, начальная популяция состоит только из «жизнеспособных» хромосом. Это же требование проверяется и в процессе реализации КМНК, «нежизнеспособным» хромосомам ставится в соответствие максимально возможное значение функции соответствия с целью их дальнейшего исключения из популяции. При реализации ГА выполняется одноточечное скрещивание. Анализ результатов кластеризации при различных способах кодирования хромосом показал, что кодирование хромосомы степенями принадлежности объектов ценграм кластеров целесообразно использовать, если количество объектов невелико и не превосходит количество критериев в 2-5 раз.

Ниже приведен КМНК при кодировании хромосомы координатами центров кластеров.

- 1. Выполняется один шаг FCM-алгоритма на основе HMT1 при формировании хромосом начальной популяции размером P.
- 2. При g < G (G и g максимальное и текущее количество поколений Γ A) выполняется один шаг Γ A с реализацией операций скрещивания и мутации и вычислением значений функции соответствия для новой популяции хромосом размером ($P + R_c \cdot P$).
- 3. Для новой популяции размером $(P+R_c\cdot P)$ выполняется один шаг FCM-алгоритма на основе HMT1 с вычислением новых значений ФП объектов центрам кластеров в соответствии с формулой (21), новых координат центров кластеров в соответствии с формулой (20) и значений функции соответствия.
- 4. Из расширенной популяции размером $(2 \cdot P + R_c \cdot P)$, полученной путем объединения популяции размером P предыдущего шага и популяции размером $(P + R_c \cdot P)$ текущего шага, удаляются «нежизнеспособные» $(P + R_c \cdot P)$ хромосом с максимальными значениями функции соответствия. Если g < G, осуществляется переход к шагу 2. Если $g \ge G$, то работа ΓA завершается и осуществляется переход к шагу 5.
- 5. Выбирается лучшая хромосома, которая минимизирует функцию соответствия. Искомые координаты центров нечетких кластеров определяются на основе лучшей хромосомы. В качестве результирующих нечетких степеней принадлежности объектов центрам кластеров полагаются степени принадлежности объектов центрам кластеров, соответствующие лучшей хромосоме, вычисленные в ходе реализации КМНК.

При реализации ГА в каждом поколении наряду с «главной» популяцией хромосом, закодированных координатами центров кластеров, вычисляются «вспомогательные» популяция степеней принадлежности объектов центрам кластеров и значений функции соответствия. Для повышения эффективности ГА предлагается в процессе формирования расширенной популяции хромосом очередного поколения ГА учитывать популяцию размером P, полученную в предыдущем поколении, и новую популяцию размером $(P + R_c \cdot P)$, полученную в текущем поколении. Для хромосом новой популяции размером $(P + R_c \cdot P)$ выполняется один шаг FCM-алгоритма на основе HMT1. Значения функции соответствия хромосом популяции размером P, полученной в предыдущем поколении, и новой популяции размером $(P + R_c \cdot P)$ текущего поколения на первых шагах КМНК могут существенно отличаться друг от друга (в новой популяции размером $(P + R_c \cdot P)$ могут присутствовать «нежизнеспособные» хромосомы с максимально возможными значениями функции соответствия). В результате формируется расширенная популяция размером $(2 \cdot P + R_c \cdot P)$, хромосомы которой упорядочиваются по возрастанию значений функции соответствия. Результирующая популяция размера Р получается путём отбрасывания из расширенной популяции размером $(2 \cdot P + R_c \cdot P)$ хромосом с худшими (максимальными) значениями функции соответствия.

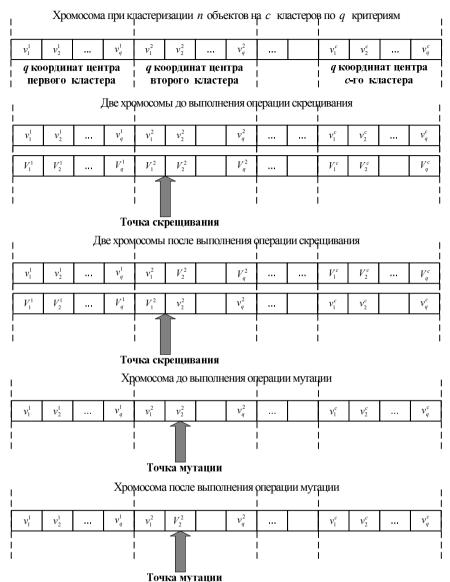


Рисунок 6 – Кодирование хромосомы координатами центров кластеров

КМНК позволяет существенно уменьшить время реализации, необходимое для получения адекватных результатов кластеризации, по сравнению со временем реализации классического FCM-алгоритма в 3-10 раз. Оптимальный размер популяции, обеспечивающий генерацию «лучшей» хромосомы (определяемой по минимальному значению функции соответствия) при кодировании хромосомы координатами центров кластеров (рисунок 6) для данного множества объектов в 96-98 случаях из 100, составляет 40-50 хромосом. Реализация ГА без использования FCM-алгоритма как самостоятельного шага вычислений оказывается менее эффективной, так как результаты кластеризации будут

существенно зависеть от того, насколько хорошо выполнена инициализация центров кластеров (а для получения адекватных результатов кластеризации требуется использование популяций значительно больших размеров).

Также в п. 4.4 рассмотрен КМНК при кодировании хромосомы степенями принадлежности объектов центрам кластеров. В п. 4.6 предложены комбинированные методы кластеризации с использованием РСМ-алгоритма на основе НМТ1, РFСМ-алгоритма на основе НМТ1, RFСМ-алгоритма на основе НМТ1 и ГА. Кодирование хромосомы осуществляется координатами центров кластеров. Инициализация данных алгоритмов осуществляется с использованием FСМ-алгоритма. Кроме того, обеспечивается расчет параметров η_j ($j=\overline{1,c}$) или δ по формулам (30) и (31) для хромосом популяции с формированием «вспомогательной» популяции для параметров η_j или δ . В результате применения этих методов выполняется минимизация целевых функций соответствующих алгоритмов кластеризации. В п. 4.7 предложен метод кластеризации объектов, представленных мультимножествами, с использованием FСМ-алгоритма на основе НМТ1 и ГА.

Предложенные методы кластеризации были использованы при решении задачи многокритериального анализа объектов жилого фонда, которая возникает, например, при составлении плана ремонтных работ, стоимостной оценке зданий и т.п.

В главе 5 «Методы и алгоритмы кластеризации на основе нечетких множеств второго типа и генетических алгоритмов» разрабатываются методы кластеризации с

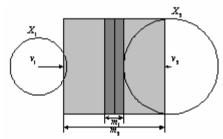


Рисунок 7—Н еопределенная максимальная нечеткая область, построенная с помощью двух фаззификагоров m_1 и m_2

у разрабатываются методы властеризации и использованием алгоритмов кластеризации на основе ИНМТ2³ и ГА, обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры существенно разной объема, с минимально возможными временными затратами

Если кластеры в множестве объектов имеют существенно разную плотность или существенно разный объем, то работа алгоритма кластеризации на основе НМТ1 существенно зависит от выбора параметров алгоритма кластеризации. FCM-алгоритм на основе НМТ1 может дать пло-

хие результаты кластеризации при выборе несоответствующего фаззификатора m из-за различия в плотности или объеме между кластерами. Идеальная ситуация — получение максимальной нечеткой области с широкой левой и узкой правой подобластями (рисунок 7). В п. 5.1 рассматривается FCM-алгоригм на основе ИНМТ2. Расширение множества объектов кластеризации на ИНМТ2 позволяет управлять неопределенностью, связанной с заданием максимальной нечеткой области за счет введения двух фаззификаторов m_1 и m_2 , и существенно улучшить результаты кластеризации. При определении интервальных первичных ФП объекта x_i рассматриваются «нижняя» и «верхняя» ФП с использованием двух различных значений фаззификатора m:

$$\overline{u}_j(x_i) = \begin{cases} u_j^1(x_i), \text{если } u_j^1(x_i) > u_j^2(x_i) \\ u_j^2(x_i), \text{если } u_j^1(x_i) \leq u_j^2(x_i) \end{cases}, \quad \underline{u}_j(x_i) = \begin{cases} u_j^1(x_i), \text{если } u_j^1(x_i) \leq u_j^2(x_i) \\ u_j^2(x_i), \text{если } u_j^1(x_i) > u_j^2(x_i) \end{cases}, \quad (32)$$

_

 $^{^3}$ Hwang C., Rhee F.C.-H. Uncertain fuzzy clustering: interval type-2 fuzzy approach to C-means // IEEE Transactions on fuzzy systems, 2007. – Vol. 15. – No 1 – P. 107-120.

где
$$u_j^p(x_i) = 1 / \sum_{k=1}^c (d_{ji} / d_{ki})^{2/(m_p - 1)}$$
 , $p = 1, 2$.

«Отпечаток неопределенности» ИНМТ2 определяется через ФП вида (32).

Использование фаззификаторов m_1 и m_2 , определяющих различные степени нечет-кости, приводит к различным целевым функциям, которые необходимо минимизировать при $m=m_1$ и $m=m_2$:

$$J_{m_1}(U,V) = \sum_{i=1}^{c} \sum_{i=1}^{n} (u_j(x_i))^{m_1} \cdot d_{ji}^2 , \qquad (33)$$

$$J_{m_2}(U,V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (u_j(x_i))^{m_2} \cdot d_{ji}^2.$$
(34)

В п. 5.2 в качестве показателя качества кластеризации для множества объектов, содержащего кластеры гиперсферической формы, предлагается использовать индекс *Sph*:

$$Sph = \sum_{j=1}^{c} r_j / \min_{\substack{j=1,c;\ t=1,c\\t\neq j}} d(v_j, v_t), (r_j = \max_{x_i \in X_j} d(x_i, v_j)),$$
(35)

а для кластеров гиперэллиптической формы – общий гиперобъем:

$$H = \sum_{i=1}^{c} (det(R_j))^{-1/2}, (R_j = \frac{1}{n_j} \cdot \sum_{i=1}^{n} (x_i - v_j) \cdot (x_i - v_j)^T),$$
(36)

где v_j — вектор координат центра j -го кластера; x_i — вектор координат i -го объекта; $d(x_i,v_j)$ — расстояние между объектом x_i , принадлежащим кластеру X_j , и центром кластера v_j ; $d(v_j,v_t)$ — расстояние между центрами кластеров v_j и v_t ; n — количество объектов; c — количество кластеров; R_j — ковариационная матрица j -го кластера; n_j — количество объектов, отнесенных к j -му кластеру; $det(R_j)$ — определитель ковариационной матрицы R_i .

Оба показателя качества кластеризации по формулам (35) и (36) должны быть минимизированы. Обычно общий гиперобьем H используется для оценки качества кластеризации на основе классического четкого алгоритма c -средних. Однако в данном случае применение показателей качества, используемых для алгоритмов кластеризации на основе HMT1, невозможно ввиду наличия, например, в FCM-алгоритме на основе ИНМТ2 двух фаззификаторов m_1 и m_2 . Анализ результатов кластеризации для различных тестовых примеров с использованием показателей качества кластеризации по формулам (35) и (36) показал целесообразность их применения на практике.

В п. 5.4 рассматривается РСМ-алгоритм на основе ИНМТ2, для которого определяются две целевые функции при $m = m_1$ и $m = m_2$:

$$J_{m_1}(W,V) = \sum_{i=1}^{c} \sum_{i=1}^{n} (w_j(x_i))^{m_1} \cdot d_{ji}^2 + \sum_{i=1}^{c} \eta_j^2 \cdot \sum_{i=1}^{n} (1 - w_j(x_i))^{m_1},$$
(37)

$$J_{m_2}(W,V) = \sum_{j=1}^{c} \sum_{i=1}^{n} (w_j(x_i))^{m_2} \cdot d_{ji}^2 + \sum_{j=1}^{c} \eta_j^2 \cdot \sum_{i=1}^{n} (1 - w_j(x_i))^{m_2} ,$$
(38)

при этом в качестве $u_j^p(x_i)$ в формулах (32) используются $\Phi T \ w_j^p(x_i)$ по формуле (25).

Управление неопределенностью фаззификатора m осуществляется с помощью вычисления центров кластеров и получения четкого разбиения для принятия конечного решения о результатах кластеризации. По принципу расширения центроид ИНМТ2 \widetilde{X} вычисляется как

$$v_{\widetilde{X}} = \sum_{u(x_i) \in J_x} \dots \sum_{u(x_r) \in J_x} \frac{1}{i=1} / \left(\sum_{i=1}^n x_i \cdot u(x_i)^m / \sum_{i=1}^n u(x_i)^m \right) = [v_{left}, v_{right}].$$
(39)

 $v_{\widetilde{X}} = \sum_{u(x_1) \in J_{x_1}} \dots \sum_{u(x_n) \in J_{x_n}} 1 / (\sum_{i=1}^n x_i \cdot u(x_i)^m / \sum_{i=1}^n u(x_i)^m) = [v_{left}, v_{right}].$ (39) При оценке центров кластеров используется итерационный алгоритм Карника — Менделя. Центры кластеров представляются интервалом $v_{\widetilde{X}} = [v_{left}, v_{right}]$, а центроид ИНМТ2 находится как $v = (v_{left} + v_{right})/2$. При этом, если сложность реализации FCM-алгоритма на основе HMT1 оценивается как $O(n^2)$, то сложность реализации FCM-алгоритма на основе ИНМТ2 – как $O(n^3)$ (за счет использования алгоритма Карника – Менделя). В пп. 5.3 и 5.4 для поиска оптимальной комбинации фаззификаторов m₁ и m₂ для FCM-алгоритма на основе ИНМТ2 и оптимальной комбинации фаззификаторов m_1 , m_2 и значений «ширины зоны» η_j ($j=\overline{1,c}$) для РСМ-алгоритма на основе ИНТМ2 предлагается использовать соответствующие ГА. При этом для FCM-алгоритма хромосома задается в виде $s = (m_1, m_2)$, а для PCM-алгоритма — в виде $s = (m_1, m_2, \eta_1, ..., \eta_c) \;, \quad \text{fig.} \quad m_1, m_2 \in (1, m_{max}] \;; \quad m_1 < m_2 \;; \quad \eta_j \in [\eta_j^{min}, \eta_j^{max}] \quad \big(\; j = \overline{1,c} \; \big);$ m_{max} , η_{i}^{min} и η_{i}^{max} — некоторые действительные числа. Если для инициализации РСМлиторитма на основе ИНМТ2 используются результаты кластеризации, полученные с помощью FCM-алгоритма на основе ИНМТ2, то хромосому целесообразно представлять в виде: $s = (\eta_1, ..., \eta_c) \ (m_1, m_2 - фиксированы).$

ГА для FCM-алгоритма на основе ИНМТ2 представляется в следующем виде.

- Случайным образом создается популяция размером P. При этом выполняется проверка условия: $m_1 < m_2$.
- 2. При g < G (G и g максимальное и текущее количество поколений ГА соответственно) реализуется FCM-алгоритм на основе ИНМТ2 с вычислением значений функции соответствия для каждой хромосомы и создается $R_c \cdot P/2$ пар хромосом-родителей.
- 3. Выполняются операции скрещивания и мугации для текущей популяции. При этом выполняется проверка условия: $m_1 < m_2$. Для хромосом-отпрысков реализуется FCMалгоритм на основе ИНМТ2 и вычисляются значения функции соответствия.
- 4. Создается новая популяция размером $(P + R_c^{-1}P)$, дополненная хромосомамиотпрысками в количестве $R_c \cdot P$, затем $R_c \cdot P$ хромосом с худшими значениями функции соответствия отбрасываются. Если g < G, осуществляется переход к шагу 2, в противном случае – переход к шагу 5.
- 5. Выбирается лучшая хромосома, которая минимизирует функцию соответствия. Для объекта x_i (i = 1, n) определяется принадлежность к кластерам.

Одновременно с «главной» популяцией хромосом вида $s = (m_1, m_2)$ существуют «вспомогательные» популяции значений функции соответствия, координат центров кластеров и степеней принадлежности объектов центрам кластеров.

В п. 5.4 предложены два метода кластеризации с использованием РСМ-алгоритма на основе ИНМТ2 и ГА: в одном случае управление неопределенностью осуществляется за счет выбора двух фаззификаторов m_1 и m_2 (как описано выше); во втором случае — за счет выбора комбинации значений «ширины зоны» η_i : η_{i1} и η_{i2} , где $\eta_{i1} < \eta_{i2}$ (j=1,c). В п. 5.5 предложен метод кластеризации с использованием FCM- и PCMалгоритмов на основе ИНМТ2 и ГА для случая, когда существует неопределенность в выборе алгоритма кластеризации (FCM-PCM-алгоритм на основе ИНМТ2). Для каждого метода кластеризации реализуется вычисление «нижней» и «верхней» ФП по формулам (32) в соответствии с реализуемым способом представления неопределенности.

Так как алгоритм кластеризации на основе ИНТМ2 обеспечивает поиск локальнооптимального разбиения, то только многократное выполнение метода кластеризации на основе ИНМТ2 при фиксированных значениях параметров алгоритма кластеризации позволяет найти адекватные результаты кластеризации.

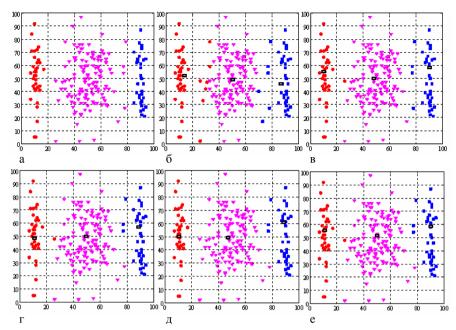


Рисунок 8 – Кластеризация объектов на основе ИНМТ2:

а – множество объектов кластеризации;

 δ – FCM-алгоритм на основе HMT1 при m=2;

в – FCM-алгоритм на основе ИНМТ2 при $m_1 = 108,813$ и $m_2 = 112,174$;

г
– РСМ-алгоритм на основе ИНМТ2 при $m_1=108,\!813$,
 $m_2=112,\!174$, $\eta_1=115,\!411$, $\eta_2=148,\!455$,
 $\eta_3=117,\!808$;

д—РСМ-алгоритм на основе ИНМТ2 при m = 120,134 , η_{11} = 89,139 , η_{12} = 104,895 , η_{13} = 79,452 , η_{21} = 89,270 , η_{22} = 117,155 , η_{23} = 93,426 ;

е – FCM-РСМ-алгоритм на основе ИНМТ2 при m = 135,794 , η_1 = 78,737 , η_2 = 122,235 , η_3 = 75,345

Таблица 3

Алгоритм	Значение функции	Ошибочно
кластеризации	соответствия	классифицир. объекты
-	(для сгенерирован. кластеров) 279,725273	0 объектов
Алгоритм четких <i>с</i> -средних	350,616421	9 объектов (4,5 %)
FCM на основе HMT1 при <i>m</i> =2	(нечетк. общий гиперобьем) 541,268379	13 объектов (6,5 %)
FCM на основе ИНМТ2	330,143524	5 объектов (2,5 %)
РСМ на основе ИНМТ2 (произвольн. значения фаззификаторов)	311,775548	4 объекта (2 %)
РСМ на основе ИНМТ2 (произвольн. значения «ширины зоны»)	306,787457	3 объекта (1,5 %)
FCM-PCM на основе ИНМТ2	319,708034	4 объекта (2 %)

В связи с этим в п. 5.6 предлагаются двухуровневые ГА для алгоритмов кластеризации на основе ИНМТ2, позволяющие решить проблему, связанную с получением локально-оптимального разбиения. При этом на первом уровне двухуровнего ГА осуществляется поиск оптимальных значений параметров алгоритма кластеризации на основе ИНМТ2, а на втором – поиск оптимального расположения центров кластеров и значений степеней принадлежности при фиксированных значениях параметров алгоритма кластеризации.

На рисунке 8, а показано множество объектов, содержащее три кластера существенно разного объема (объекты разных кластеров помечены маркерами разной формы и разного цвета). Кластеры представляют собой множества объектов, координаты которых были сгенерированы с использованием нормального закона распределения с центрами (10, 50). (50, 50) и (90, 50). При этом при генерации первой и второй координат объектов первого и третьего кластеров использовались нормальные законы распределения, имеющие одинаковые дисперсии. Координаты объектов второго кластера имеют существенно большую дисперсию по обеим координатам. Первый и второй кластеры содержат по 35 объектов, а второй кластер – 130 объектов. На рисунке 8, б-е приведены результаты кластеризации с использованием FCM-алгоритма на основе НМТ1 и алгоритмов кластеризации на основе ИНМТ2, оптимальные параметры которых были определены с помощью ГА, демонстрирующие эффективность применения ИНМТ2 (центры кластеров помечены маркерами-«прямоугольниками»). В таблице 3 приведены результаты анализа качества кластеризации с использованием указанных выше методов, демонстрирующие уменьшение количества ошибочно классифицированных (распознанных) объектов при применении ИНМТ2.

Методы кластеризации, реализующие совместное использование алгоритмов кластеризации на основе ИНМТ2 и ГА, позволяют адекватно управлять неопределенностью, которая возникает при анализе множества объектов, образованного из кластеров существенно разной плотности или существенно разного объема. На тестовых примерах доказана целесообразность использования предлагаемых показателей качества кластеризации. Применение ГА позволяет найти оптимальную комбинацию параметров алгоритма кластеризации на основе ИНМТ2, обеспечивающую адекватные результаты кластеризации, что подтверждается минимальным значением функции соответствия.

Методы кластеризации с применением алгоритмов кластеризации на основе ИНМТ2 и ГА были использованы при выполнении многокритериального анализа технического состояния зданий и сооружений.

Глава 6 «Программная реализация методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности» посвящена применению разработанных в диссертации методов, моделей и алгоритмов для решения задач принятия решений в условиях неопределенности, сформулированных в главе 1.

Дано краткое описание разработанных пакетов прикладных программ, реализованных в среде MATLAB 7.0, приведены их структурные схемы. ППП «Модели прогнозирования на основе нечетких множеств первого и второго типов», обеспечивающий решение задач краткосрочного прогнозирования (на 1 шаг вперед) процессов с преобладающей детерминированной составляющей, представленных ВР с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса, разработан с использованием нечетких моделей прогнозирования, описанных в главе 2. ППП «Нечеткие городские инженерные коммуникации», предназначенный для комплексной диагностики состояния ГИК: принятия решения о причинах неисправностей ГИК и предпосылок неисправностей ГИК при необходимости учёта территориальной распределённости объектов на основе алгоритмов идентификации неисправностей и предпосылок неисправностей; принятия решения о возможности перекрытия того или иного вентиля на аварийном участке при минимально возможной численности населения, временно отключаемого от подачи воды на основе иерархической системы нечеткого вывода, реализован с использованием алгоритмов и моделей, описанных в главе 3. Программная реализация ППП «Упорядочение и классификация инвестиционных проектов на основе мультимножеств», предназначенного для комплексного многокритериального анализа ИП при ярко выраженном наличии несовпадающих, в том числе противоречивых, оценок экспертов: принятия решения о классификации ИП на основе мультимножеств и индивидуальных систем нечеткого вывода, настроенных с помощью ГА; принятия решения об упорядочении ИП на основе парных сравнений по шкале Саати, нечеткого метода Дельфы и мультимножеств; принятия решения о кластеризации ИП с использованием FCM-алгоритма на основе НМТ1 и ГА, выполнена с использованием методов упорядочения, классификации и кластеризации объектов, представленных мультимножествами, изложенных в главах 3 и 4. ППП «Настройка систем нечеткого вывода классификации способов несанкционированного отбора электроэнергии с использованием генетических алгоритмов», обеспечивающий выполнение настройки параметров систем нечеткого вывода абонентов электросети с использованием их индивидуальных моделей энергопотребления при наличии графического и/или аналитического представления данных обучающей выборки, реализован с использованием алгоритмов, предложенных в главе 3. Программная реализация ППП «Оценка обобщенного мнения экспертов на основе нечеткого метода Дельфы и центроидов нечетких множеств первого и второго типов», предназначенного для решения задачи оценки потенциального качества изменяемых многомерных объектов при малых объежах априорной информации и формирования обобщенного мнения экспертов на основе нечеткого метода Дельфы с использованием центроидов НМТ1, а в сложных случаях – центроидов ИНМТ2, применение которых позволяет дифференцировать объекты, имеющие равные центроиды НМТ1 для оценок по некоторым критериям, выполнена с использованием методов оценивания объектов на основе НМТ1 и ИНМТ2, изложенных в главе 3. ППП «Методы кластеризации объектов на основе нечетких множеств первого и второго типов», обеспечивающий выполнение многокритериальной кластеризации совокупностей объектов (при нечётком определении состояний многомерных объектов), содержащих как кластеры подобной плотности, подобного объема и подобной мощности, так и кластеры существенно разной плотности, существенно разного объема или существенно разной мощности, разработан с использованием методов и алгоритмов, изложенных в главах 4 и 5.

В заключении диссертационной работы сформулированы основные научные результаты, полученные в рамках решения поставленной научно-технической проблемы разработки эффективных методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности на основе комплексного использования инструментария теории нечётких множеств и генетических алгоритмов.

- 1. Выполнено исследование проблемы поддержки принятия решений в условиях неопределенности, выявлены достоинства и недостатки известных методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности. Проведенный анализ показал, что существующие методы, модели и алгоритмы поддержки принятия решений в условиях неопределенности зачастую не обеспечивают принятие объективных и адекватных решений ввиду недостаточно обоснованного выбора параметров моделирования, а поиск эффективных решений сопровождается значительными временными затратами из-за необходимости выполнения многократных реализаций используемых методов, моделей и алгоритмов с целью выбора оптимальных параметров моделирования. Сделан вывод о целесообразности использования генетических алгоритмов для решения задачи поиска оптимальных параметров методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности.
- 2. Разработаны модели краткосрочного прогнозирования процессов с преобладающей детерминированной составляющей, представленных ВР с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса, на основе НМТ1 и ИНМТ2 с использованием генетических алгоритмов, обеспечивающие повышение точности прогнозирования за счет выбора оптимальных параметров моделей прогнозирования.

- 3. Разработаны генетические алгоритмы «тонкой» настройки параметров систем нечеткого вывода на основе HMT1 и ИНМТ2 с использованием обучающих выборок, обеспечивающие повышение объективности и адекватности принимаемых решений с применением моделей поддержки принятия решений в условиях неопределенности.
- 4. Разработаны методы упорядочения, классификации и кластеризации объектов при наличии несовпадающих, в том числе противоречивых, исходных данных с использованием мультимножеств, нечеткого метода Дельфы, схемы Беллмана Заде и систем нечеткого вывода, позволяющие учесть в процессе принятия решения все, в том числе и противоречивые, оценки объектов без использования дополнительных преобразований типа усреднения и смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных. Предложено при формировании обобщающих правил классификации объектов, представленных мультимножествами, для выполнения предварительной сортировки объектов применять индивидуальные системы нечеткого вывода, параметры которых настроены с использованием генетических алгоритмов.
- Предложен метод оценивания объектов с вычислением центроидов ИНМТ2 на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы, позволяющий дифференцировать объекты, имеющие одинаковые центроиды НМТ1 на основе этих же экспертных оценок.
- 6. Разработаны методы кластеризации объектов с использованием модификаций алгоритма нечетких c -средних на основе HMT1 и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры подобной плотности и подобного объема, с минимальными временными затратами.
- 7. Разработаны методы кластеризации объектов с использованием модификаций алгоритма нечетких c-средних на основе ИНМТ2 и генетических алгоритмов, позволяющие учесть свойства кластерной типичности и кластерной относительности и обеспечивающие получение адекватных результатов кластеризации множества объектов, содержащего кластеры существенно разной плотности или существенно разного объема, с минимальными временными затратами.
- 8. Разработаны показатели качества кластеризации, позволяющие получить адекватные результаты кластеризации множества объектов, содержащего кластеры гиперсферической или гиперэллипсоидной формы, с использованием методов кластеризации на основе ИНМТ2.
- Разработаны пакеты прикладных программ, реализующие предлагаемые методы, модели и алгоритмы поддержки принятия решений в условиях неопределенности, и решен ряд актуальных прикладных задач:
- задача прогнозирования процессов с преобладающей детерминированной составляющей, представленных временными рядами с короткой длиной актуальной части, в условиях отсутствия априорной информации о вероятностных характеристиках процесса;
- задача принятия инвестиционных решений на основе многокритериального упорядочения, классификации и кластеризации ИП при ярко выраженном наличии несовпадающих, в том числе противоречивых, оценок экспертов;
- задача диагностики состояния городских инженерных коммуникаций и принятия решений в штатных и аварийных ситуациях при необходимости учёта территориальной распределённости объектов;
- задача классификации способов несанкционированного отбора электроэнергии при наличии графического и/или аналитического представления данных обучающей выборки;
- задача оценки недвижимости как задача оценки потенциального качества изменяемых многомерных объектов при малых объежах априорной информации;

 задача технического контроля состояния зданий и сооружений как задача кластеризации при нечётком определении состояний многомерных объектов.

Неопределенность в данных задачах связана с невозможностью полного сбора и учёта информации о влияющих воздействиях, с неточностью выполняемых измерений, а так же с неопределенностью, неполнотой и нечеткостью знаний экспертов в конкретной предметной области, то есть с проявлением субъективного человеческого фактора.

- 10. Исследования разработанных методов, моделей и алгоритмов показали:
- повышение обоснованности и адекватности принятия решения в условиях неопределенности и неточности исходной информации, в том числе при несовпадающих (противоречивых) оценках экспертов;
- повышение качества формализации опыта экспертов, который зачастую является единственной наиболее достоверной информацией при решении многих задач поддержки принятия решений;
- снижение временных и финансовых затрат, связанных как с необходимостью сбора и учета точных и полных исходных данных (что может быть принципиально невозможным), так и с разработкой сложных классических математических моделей или необходимостью многократной реализации классических методов, моделей и алгоритмов с целью выбора соответствующих оптимальных параметров, обеспечивающих принятие адекватных решений.

Анализ приведенных результатов дает основание полагать, что данная работа связана с решением важной научно-технической проблемы и представляет собой развитие методов, моделей и алгоритмов поддержки принятия решений в условиях неопределенности на основе комплексного использования инструментария теории нечётких множеств и генетических алгоритмов.

В приложении 1 приведены акты внедрения результатов диссертационной работы и свидетельства о регистрации программ в ФГУ ФИПС (РОСПАТЕНТ) и ОФАП. В приложениях 2-5, включенных в отдельный том, содержатся дополнительные материалы к главам 2-5 (номер приложения совпадает с номером главы, к которой оно относится). В приложении 2 приведены примеры прогнозирования тенденций рынка труда, метеоданных и показателей, связанных с внешнеторговой деятельностью, с использованием моделей прогнозирования на основе НМТ1 и ИНМТ2; выполнен сравнительный анализ с результатами прогнозирования на основе классических моделей (модели Бурга, авторегрессионной модели и полиномиальной модели). В приложении 3 приведены примеры решения задачи оценки влияния изменений состояния окружающей среды на стоимость недвижимости с использованием центроидов НМТ1 и ИНМТ2; задачи диагностики состояния ГИК и принятия решений в штатных и аварийных ситуациях при необходимости учёта территориальной распределённости объектов; задачи классификации способов несанкционированного отбора электроэнергии с использованием систем нечеткого вывода на основе НМТ1 и ИНМТ2; задачи принятия инвестиционных решений на основе многокритериального упорядочения, классификации и кластеризации ИП, представленных мультимножествами с использованием индивидуальных систем нечеткого вывода для предварительной экспертной сортировки ИП. В приложении 4 приведены примеры кластеризации с использованием алгоритмов кластеризации на основе НМТ1, в том числе примеры кластеризации при оценке технического состояния зданий и сооружений. В приложении 5 приведены примеры кластеризации с использованием алгоритмов кластеризации на основе ИНМТ2; выполнен сравнительный анализ с результатами кластеризации с использованием алгоритмов кластеризации на основе НМТ1. В приложениях выполнена оценка затрат памяти и сложности реализации предлагаемых моделей и методов.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи в центральных журналах перечня ВАК

- 1. Демидова, Л.А. Идентификация неисправностей в городских инженерных коммуникациях на основе обратного нечеткого вывода [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Вестник РГРТА. Рязань, 2005. Вып. 16. С. 3-10.
- 2. Демидова, Л.А. Иерархическая система нечеткого вывода для диагностики состояния городских инженерных коммуникаций [Текст] / Л.А. Демидова, В.В. Кираковский // Системы управления и информационные технологии. Воронеж, 2005. № 5 (22). С. 36-42.
- 3. Демидова, Л.А. Диагностика городских инженерных коммуникаций на основе иерархической системы нечеткого вывода [Текст] / Л.А. Демидова, В.В. Кираковский // Вестник РГРТА. Рязань, 2005. Вып. 17. С. 57-65.
- 4. Демидова, Л.А. Генерирование решающих правил классификации инвестиционных проектов на основе систем нечеткого вывода и мультимножеств [Текст] / М.В. Гусева, Л.А. Демидова // Системы управления и информационные технологии. Воронеж, 2006. № 4 (26). С. 46-53.
- 5. Демидова, Л.А. Классификация инвестиционных проектов на основе систем нечеткого вывода и мультимножеств [Текст] / М.В. Гусева, Л.А. Демидова // Вестник РГРТА. Рязань, 2006. Вып. 19. С. 157-166.
- 6. Демидова, Л.А. Многокритериальная классификация инвестиционных проектов на основе систем нечеткого вывода и мультимножеств [Текст] / М.В. Гусева, Л.А. Демидова // Научнотехническая информация. Информационные процессы и системы, 2006. —№12. С. 16-20.
- 7. Демидова, Л.А. Классификация объектов на основе мультимножеств и нечеткой кластеризации [Текст] / Л.А. Демидова // Известия ТРТУ. Таганрог, 2006. №15(70). С. 72-79.
- 8. Демидова, Л.А. Многокритериальная классификация объектов на основе мультимножеств и нечеткой кластеризации [Текст] / Л.А. Демидова // Известия ТулГУ. Серия «Радиотехника и радиооптика». Тула: Изд-во ТулГУ, 2006. Выпуск 1. Т. 8. С. 45-56.
- 9. Демидова, Л.А. Генерирование правил классификации участков сетей городских инженерных коммуникаций [Текст] / Л.А. Демидова // Известия ТулГУ. Серия «Вычислительная техника. Информационные технологии. Системы управления». Тула: Изд-во ТулГУ, 2006. Выпуск 3. Т. 1.— С. 151-158.
- 10. Демидова, Л.А. Многокритериальный анализ альтернатив на основе схемы Беллмана Заде и мультимножеств [Текст] / Л.А. Демидова, В.В. Кираковский // Известия вузов. Северо-Кавказский регион. Технические науки. Спецвыпуск «Математическое моделирование и компьютерные технологии», 2006. − С. 9-14.
- 11. Демидова, Л.А. Прогнозирование тенденций рынка труда на основе однофакторных нечетких временных рядов [Текст] / Л.А. Демидова // Системы управления и информационные технологии. Воронеж, 2007. —№ 3.2(29). С. 241-246.
- 12. Демидова, Л.А. Разработка однофакторных нечетких моделей для анализа тенденций временных рядов с использованием генетического алгоритма [Текст] / Л.А. Демидова // Научнотехнические ведомости Санкт-Петербургского государственного политехнического университета. СПб., 2007. № 4. С. 156-164.
- 13. Демидова, Л.А. Генетический алгоритм поиска оптимальных параметров однофакторной модели прогнозирования на основе непрерывных нечетких множеств типа-2 [Текст] / Л.А. Демидова // Системы управления и информационные технологии. – Воронеж, 2008. – № 2(32). – С. 43-48.
- 14. Демидова, Л.А. Генетический алгоритм поиска параметров однофакторной модели прогнозирования высокого порядка на основе нечетких множеств второго типа [Текст] / Л.А. Демидова // Проблемы полиграфии и издательского дела, 2008. № 4. С. 31-44.
- 15. Демидова, Л.А. Кластеризация объектов с использованием FCM-алгоритма на основе нечетких множеств второго типа и генетического алгоритма [Текст] / Л.А. Демидова, Е.И. Коняева // Вестник РГРТУ. Рязань, 2008. № 4 (выпуск 26). С. 46-54.
- 16. Демидова, Л.А. Оценка технического состояния зданий и сооружений на основе инструментария теории нечетких множеств [Текст] / Л.А. Демидова, В.В. Кираковский // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. СПб., 2008. № 2. С. 163-169.
- 17. Демидова, Л.А. Методы кластеризации объектов на основе нечетких множеств второго типа и генетического алгоритма [Текст] / Л.А. Демидова, В.В. Кираковский // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. СПб., 2008. № 6(69). С. 136-142.

Монографии

- Демидова, Л.А. Алгоритмы и системы нечеткого вывода при решении задач диагностики городских инженерных коммуникаций в среде МАТLАВ [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин. – М.: Радио и связь, Горячая линия – Телеком, 2005. – 365 с.: ил.
- Демидова, Л.А. Методы и алгоритмы принятия решений в задачах многокритериального анализа [Текст] / Л.А. Демидова, А.Н. Пылькин. М.: Горячая линия Телеком, 2007. 232 с.: ил.

Свидетельства о регистрации программы в ФГУ ФИПС

- 20. Демидова, Л.А. Многокритериальный анализ объектов на основе нечеткой логики и мультимножеств (FuzzyLogicMultistAnalisys) [Текст] / М.В. Гусева, Л.А. Демидова. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2007611203 от 21.03.2007.
- 21. Демидова, Л.А. Нечеткие городские инженерные коммуникации (FuzzyCityEngineeringCommunications) [Текст] / Л.А. Демидова, В.В. Кираковский. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2007611108 от 16.03.07.
- 22. Демидова, Л.А. Многокритериальное упорядочение объектов на основе парного сравнения, нечеткого метода Дельфы и мультимножеств (FuzzyEstimationAndMultisetOrdering) [Текст] / Л.А. Демидова, В.В. Кираковский. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2007611109 от 16.03.07.
- 23. Демидова, Л.А. Кластеризация объектов на основе алгоритма нечетких K-средних с использованием нечетких множеств первого типа и генетического алгоритма (FCMGeneticAlgorithmT1FS) [Текст] / Л.А. Демидова, В.В. Кираковский. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2008615534 от 19.11.2008.
- 24. Демидова, Л.А. Кластеризация объектов на основе алгоритма нечетких К-средних с использованием нечетких множеств второго типа и генетического алгоритма (FCMGeneticAlgorithmT2FS) [Текст] / Л.А. Демидова, В.В. Кираковский. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2008615535 от 19.11.2008.
- 25. Демидова, Л.А. Оценка обобщенного мнения экспертов на основе нечеткого метода Дельфы и центроидов нечетких множеств первого и второго типа (ExpertEstimationFuzzyDelfiT1T2FS) [Текст] / Л.А. Демидова, В.В. Кираковский. Свидетельство об официальной регистрации программы для ЭВМ в ФГУ ФИПС № 2008615536 от 19.11.2008.

Свидетельства о регистрации программы в ОФАП

- 26. Демидова, Л.А. Идентификация неисправностей и предпосылок неисправностей в городских инженерных коммуникациях на основе обратного нечеткого вывода [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин. Свидетельство об официальной регистрации программы для ЭВМ в ОФАП № 4717 от 6.05.2005.
- 27. Демидова, Л.А. Двухкаскадная система нечеткого вывода для принятия решения о возможности перекрытии вентиля на период ремонтных работ в городских инженерных коммуникациях [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин. Свидетельство об официальной регистрации программы для ЭВМ в ОФАП № 4718 от 6.05.2005.
- 28. Демидова, Л.А. Нечеткий многокритериальный анализ альтернатив на основе схемы Беллмана — Заде и мультимножеств [Текст] / М.В. Гусева, Л.А. Демидова. Свидетельство об официальной регистрации программы для ЭВМ в ОФАП № 5976 от 12.04.2006.
- 29. Демидова, Л.А. Классификация инвестиционных проектов на основе систем нечеткого вывода, мультимножеств и генетических алгоритмов [Текст] / М.В. Гусева, Л.А. Демидова. Свидетельство об официальной регистрации программы для ЭВМ в ОФАП №7449 от 27.12.2006.
- 30. Демидова, Л.А. Многокритериальная классификация на основе мультимножеств и нечеткой кластеризации [Текст] / Л.А. Демидова. Свидетельство об официальной регистрации программы для ЭВМ в ОФАП № 6264 от 26.05.2006.
- 31. Демидова, Л.А. Комплект программ для кластеризации строительных сооружений с помощью комбинирования генетического и FCM-алгоритмов [Текст] / Л.А. Демидова, Е.И. Коняева. Свидетельство об отраслевой регистрации разработки в ОФАП № 11502 от 08.09.2008.
- 32. Демидова, Л.А. Комплекс программ кластеризации объектов с использованием FCMалгоритма на основе нечетких множеств второго типа и генетического алгоритма [Текст] / Л.А. Демидова, Е.И. Коняева. Свидетельство об отраслевой регистрации разработки в ОФАП № 11549 от 24.09.2008.

- 33. Демидова, Л.А. Реализация двухуровневого генетического алгоритма кластеризации объектов на основе FCM-алгоритма с использованием нечётких множеств второго типа [Текст] / Л.А. Демидова, А.Н. Коротаев. Свидетельство об отраслевой регистрации разработки в ОФАП № 11813 от 20.11.2008.
- 34. Демидова, Л.А. Комплекс программ формирования обобщенного мнения экспертов на основе центроидов интервальных непрерывных нечетких множеств второго типа [Текст] / Л.А. Демидова, С.Б. Титов. Свидетельство об отраслевой регистрации разработки в ОФАП № 10623 от 12.05.2008.

Статьи в научно-технических журналах

- 35. Демидова, Л.А. Двухкаскадная система нечеткого вывода для решения задачи диагностики состояния городских инженерных коммуникаций [Текст] / Л.А. Демидова, В.В. Кираковский // Информационные технологии моделирования и управления. Воронеж: Издательство «Научная книга», 2005. —№ 1(19). С. 9-20.
- 36. Демидова, Л.А. Технология идентификации неисправностей в городских инженерных коммуникациях на основе обратного нечеткого вывода [Текст] / Л.А. Демидова, А. Н. Пылькин // Информация и космос. СПб., 2005. № 4. С. 42-47.
- 37. Демидова, Л.А. Technology of malfunction identification in city engineering communications on the base of the return fuzzy conclusion [Текст] / Л.А. Демидова, А. Н. Пылькин // European journal of natural history. London-Moscow, 2006. № 1. С. 82-87.
- 38. Демидова, Л.А. Многокритериальное упорядочение участков сетей городских инженерных коммуникаций по противоречивым данным на основе нечеткого метода Дельфы и теории мультимножеств [Текст] / Л.А. Демидова // Научно-техническая информация. Серия 2. Информационные процессы и системы, 2007. № 11. С. 20-25.
- 39. Демидова, Л.А. Прогнозирование тенденций рынка труда на основе однофакторных нечетких временных рядов и генетического алгоритма [Текст] / Л.А. Демидова // Вестник РГРТУ. Рязань, 2008. № 2 (выпуск 24). С. 75-82.
- 40. Демидова, Л.А. Прогнозирование тенденций временных рядов на основе однофакторной нечеткой модели с использованием дискретных нечетких множеств второго типа и генетического алгоритма [Текст] / Л.А. Демидова // Бизнес-информатика, 2008. № 4(6). С. 46-53.
- 41. Демидова, Л.А. Генетический алгоритм поиска оптимальной комбинации фаззификаторов для FCM-алгоритма на основе нечетких множеств второго типа [Текст] / Л.А. Демидова, Е.И. Коняева, А.Н. Коротаев // Информационные технологии моделирования и управления. Воронеж: Научная книга, 2008. № 6(49). С. 657-665.

Статьи в межвузовских сборниках

- 42. Демидова, Л.А. Исследование модели диагностики неисправностей в городских инженерных сетях [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Информационные технологии в процессе подготовки современного специалиста: межвузовский сборник. Выпуск 7. Липецк: ГОУ ВПО «ЛГПУ», 2004. С. 24-37.
- 43. Демидова, Л.А. Разработка модели диагностики неисправностей в городских инженерных сетях [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. Л.П. Коричнева. М.: Минобразования России; Рязань: РГРТА, 2004. С. 25-30.
- 44. Демидова, Л.А. Технология идентификации неисправностей в городских инженерных коммуникациях на основе обратного нечеткого вывода [Текст] / Л.А. Демидова, В.В. Кираковский // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Минобразования России; Рязань: РГРТА, 2005. С. 4-14.
- 45. Демидова, Л.А. Методы и алгоритмы системы поддержки принятия решений в многокритериальной среде на основе схемы Беллмана Заде и теории мультимножеств [Текст] / М.В. Гусева, Л.А. Демидова, А.Н. Пылькин // Математическое и программное обеспечение информационных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия-Телеком, 2006. С. 8-15.
- 46. Демидова, Л.А. Классификация объектов на основе мультимножеств и нечеткой кластеризации [Текст] / Л.А. Демидова, Р.Е. Медведев // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия-Телеком, 2006. С. 62-69.
- Демидова, Л.А. Настройка параметров функций принадлежности и весовых коэффициентов базы правил нечеткого контроллера на основе генетического алгоритма [Текст] / Л.А. Демидова, Р.Е.

Медведев, В.А. Сорокин // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. – М.: Горячая линия – Телеком, 2007. – С. 20-25.

- 48. Демидова, Л.А. Исследование проблемы согласования результатов при определении рыночной стоимости объекта оценки. Применение аппарата теории мультимножеств в рамках оценочной деятельности [Текст] / Л.А. Демидова, С.Б. Титов // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия Телеком. 2007. С. 40-44.
- 49. Демидова, Л.А. Пример реализации процедуры иерархической кластеризации объектов, представленных мультимножествами [Текст] / Л.А. Демидова, Е.И. Коняева // Математическое и программное обеспечение информационных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия Телеком, 2007. С. 34-38.
- 50. Демидова, Л.А. Классификация объектов жилого фонда на основе FCM-алгоритма и генетического алгоритма [Текст] / Л.А. Демидова, В.В. Кираковский, Е.И. Коняева // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия Телеком. 2008. С. 21-32.
- 51. Демидова, Л.А. Система поддержки принятия решений при классификации способов несанкционированного отбора электроэнергии [Текст] / Л.А. Демидова, И.Ю. Москвитин // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия Телеком, 2008. С. 47-51.
- 52. Демидова, Л.А. Классификация объектов жилой недвижимости с использованием алгоритма нечетких *с* -средних на основе нечетких множеств второго типа [Текст] / Л.А. Демидова, С.Б. Тигов // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. М.: Горячая линия Телеком, 2008. С.129-135.
- 53. Демидова, Л.А. Методы, модели и алгоритмы поддержки принятия решений в условиях неопределенности на основе теории нечетких множеств, мультимножеств и генетических алгоритмов [Текст] / Л.А. Демидова // Задачи системного анализа, управления и обработки информации: межвуз. сб. науч. тр. Вып. 2. М.: МГУП, 2008. 171 с.: ил.

Доклады на Международных и Всероссийских конференциях

- 54. Демидова, Л.А. Оценка текущего состояния узлов и элементов городских инженерных коммуникаций на основе правил нечеткого условного вывода [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Микроэлектроника и информатика. 11-я Всероссийская межвузовская научно-техническая конференция студентов и аспирантов: тезисы докладов. М.: МИЭТ. 2004. С. 255.
- 55. Демидова, Л.А. Разработка базы нечетких правил диагностики узлов городских инженерных коммуникаций [Текст] / Л.А. Демидова, В.В. Кираковский // Новые информационные технологии. Разработка и аспекты применения: материалы седьмой Всероссийской с Международным участием научной конференции молодых ученых и аспирантов. Таганрог: ТГРУ, 2004. С. 17-19.
- 56. Демидова, Л.А. Разработка модели диагностики неисправностей в городских инженерных сетях с использованием нейро-нечеткого вывода [Текст] / Л.А. Демидова, В.В. Кираковский // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 13-й Международной научно-технической конференции. Рязань: РГРТА, 2004. С. 219-220.
- 57. Демидова, Л.А. Модель диагностики неисправностей в городских инженерных сетях [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Системный анализ в проектировании и управлении: труды VIII Международной научно-практической конференции. Часть 2. СПб.: Издательство «Нестор», 2004. С. 103-105.
- 58. Демидова, Л.А. Учет текущего состояния узлов и элементов инженерных коммуникаций в ГИС [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 12-й Международной научнотехнической конференции. Рязань: РГРТА, 2004. С. 3-4.
- 59. Демидова, Л.А. Комплекс программ нечеткой диагностики городских инженерных коммуникаций [Текст] / Л.А. Демидова, В.В. Кираковский // Современные проблемы информатизации в технике и технологиях: сборник трудов X Международной открытой научной конференции. Вып. 10 / под ред. д-ра техн. наук, проф. О.Я. Кравца. Воронеж: Научная книга, 2005. С. 155-156.
- 60. Демидова, Л.А. Алгоритмы и системы нечеткого вывода при решении задач диагностики городских инженерных коммуникаций в среде MATLAB [Текст] / Л.А. Демидова, В.В. Кираковский, А.Н. Пылькин // Труды Международного форума по проблемам науки, техники и образования / под ред. В.П. Савиных, В.В. Вишневского. М., 2005. Том 3. С. 24-26.

- 61. Демидова, Л.А. Настройка нечеткой модели типа Мамдани на основе оптимизации целевой функции при решении задач диагностики городских инженерных коммуникаций [Текст] / Л.А. Демидова, А.А. Четвертаков // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 14-й Международной научно-технической конференции. Рязань: РГРТА, 2005. С. 186-187.
- 62. Демидова, Л.А. Настройка нечеткой модели типа Мамдани на основе генетических алгоритмов при решении задач диагностики городских инженерных коммуникаций [Текст] / Л.А. Демидова, А.А. Четвертаков // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 14-й Международной научно-технической конференции. Рязань: РГРТА, 2005. С. 187-189.
- 63. Демидова, Л.А. Разработка систем нечеткого вывода для классификации инвестиционных проектов [Текст] / Л.А. Демидова, М.В. Гусева // Искусственный интеллект в XXI веке. Решения в условиях неопределенности: материалы IV Международной научно-технической конференции. Пенза, 2006. С. 124—127.
- 64. Демидова, Л.А. Многокритериальная диагностика городских инженерных коммуникаций на основе данных инженерно-геодезических изысканий [Текст] / Л.А. Демидова, В.В. Кираковский // Перспективы развития инженерных изысканий в строительстве в Российской Федерации: материалы 1-й Всероссийской конференции изыскательских организаций. Часть II. М., 2006. С. 63-69.
- 65. Демидова, Л.А. Оценка влияния изменений состояния окружающей среды на стоимость недвижимости на основе нечетких множеств второго типа [Текст] / Л.А. Демидова, В.В. Кираковский // Проблемы информатики в образовании, управлении, экономике и технике: сборник статей VII Всероссийской научно-технической конференции. Пенза: Приволжский дом знаний, 2007. С. 148-150.
- 66. Демидова, Л.А. Оценка качества строительных работ на основе нечеткого метода Дельфы и нечетких множеств второго типа [Текст] / Л.А. Демидова, В.В. Кираковский // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 15-й Международной научно-технической конференции. Часть 1. Рязань: РГРТУ, 2008. С. 7.
- 67. Демидова, Л.А. Применение алгоритма нечеткой кластеризации на основе нечетких множеств второго типа к задаче анализа сложных многомерных данных [Текст] / Л.А. Демидова, П.А. Кудрин // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: материалы 15-й Международной научно-технической конференции. Часть 1. Рязань: РГРТУ, 2008. С. 41.
- 68. Демидова, Л.А. Система поддержки принятия решений на основе многокритериального анализа [Текст] / Л.А. Демидова // Новые информационные технологии в научных исследованиях и образовании: материалы XIII Всероссийской научно-технической конференции студентов. Часть II. Рязань: РГРТУ, 2008. С.135-136.
- 69. Демидова, Л.А. Оценка технического состояния строительных объектов на основе нечеткого метода Дельфы и нечетких множеств первого и второго типа [Текст] / Л.А. Демидова, В.В. Кираковский // Управление созданием и развитием систем, сетей и усгройств телекоммуникаций / под ред. дра экон. наук, канд. техн. наук, проф. А.В. Бабкина, дра техн. наук, проф. В.А. Кежаева: труды научно-практической конференции. СПб., 2008. С. 280-288.
- 70. Демидова, Л.А. Кластеризация объектов на основе нечетких множеств второго типа и генетического алгоритма [Текст] / Л.А. Демидова, В.В. Кираковский // Управление созданием и развитием систем, сетей и устройств телекоммуникаций / под ред. д-ра экон. наук, канд. техн. наук, проф. А.В. Бабкина, д-ра техн. наук, проф. В.А. Кежаева: труды Международной конференции. СПб., 2008. С. 212-222.
- 71. Демидова, Л.А. Разработка и настройка системы нечеткого вывода для классификации способа отбора электроэнергии на основе генетического алгоритма [Текст] / Л.А. Демидова, И.Ю. Москвитин // Сети, системы связи и телекоммуникации. Деятельность ВУЗа при переходе на Федеральный государственный образовательный стандарт 3-го поколения: материалы 33-й Всероссийской научно-техн. конф. В 2 ч. Ч. 1 / РВВКУС, Рязань, 2008. С. 31-33.
- 72. Демидова, Л.А. Обработка изображений на основе алгоритма нечетких к-средних с использованием нечетких множеств второго типа [Текст] / Л.А. Демидова, Р.В. Тишкин // Сети, системы связи и телекоммуникации. Деятельность ВУЗа при переходе на Федеральный государственный образовательный стандарт 3-го поколения; материалы 33-й Всероссийской научно-техн. конф. В 2 ч. Ч. 1. / РВВКУС, Рязань, 2008. С. 75-76.
- 73. Демидова, Л.А. Генетический алгоритм настройки параметров системы нечеткого вывода на основе нечетких множеств второго типа [Текст] / Л.А. Демидова, А.Н. Коротаев // Проблемы ин-

форматики в образовании, управлении, экономике и технике: сборник статей VIII Всероссийской научно-технической конференции. – Пенза: Приволжский Дом знаний, 2008. – С. 79-81.

- 74. Демидова, Л.А. Прогнозирование экономических показателей на основе нечетких множеств и генетического алгоритма [Текст] / Л.А. Демидова, А.А. Соколов, Ю.А. Токмаков // Проблемы информатики в образовании, управлении, экономике и технике: сборник статей VIII Всероссийской научно-технической конференции. Пенза: Приволжский Дом знаний, 2008. С. 256-263.
- 75. Демидова, Л.А. Генетический алгоритм поиска оптимальной комбинации значений фаззификаторов для FCM-алгоритма на основе нечетких множеств второго типа [Текст] / Л.А. Демидова, Е.И. Коняева // Искусственный интеллект в XXI веке. Решения в условиях неопределенности: сборник статей VI Всероссийской научно-технической конференции. Пенза: Приволжский Дом знаний, 2008. С. 28-30.
- 76. Демидова, Л.А. Двухуровневый генетический алгоритм кластеризации объектов на основе FCM-алгоритма с использованием нечетких множеств второго типа [Текст] / Л.А. Демидова, А.Н. Коротаев // Искусственный интеллект в XXI веке. Решения в условиях неопределенности: сборник статей VI Всероссийской научно-технической конференции. Пенза: Приволжский Дом знаний, 2008. С. 25-27.
- 77. Демидова, Л.А. Сравнительный анализ методов кластеризации на основе нечетких множеств первого и второго типа [Текст] / Л.А. Демидова, Е.И. Коняева, А.Н. Коротаев // Современные проблемы информатизации в анализе и синтезе технологических и программнотелекоммуникационных систем: сборник трудов. Вып. 14 / под ред. д-ра техн. наук, проф. О.Я. Кравца. Воронеж: Научная книга, 2009. С. 296-302.

Демидова Лилия Анатольевна

РАЗВИТИЕ МЕТОДОВ ТЕОРИИ НЕЧЁТКИХ МНОЖЕСТВ И ГЕНЕТИЧЕСКИХ АЛГОРИТМОВ ДЛЯ ЗАДАЧ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ В УСЛОВИЯХ НЕОПРЕДЕЛЁННОСТИ

А в т о р е ф е р а т диссертации на соискание ученой степени доктора технических наук

Подписано в печать . Формат бумаги 60×84 1/16. Бумага офсетная. Печать офсетная. Усл. печ. л. 2,5. Уч.-изд. л. 2,5. Тираж 100 экз. Бесплатно. Рязанский государственный радиотехнический университет. 390005, г. Рязань, ул. Гагарина, д. 59/1. Редакционно-издательский центр РГРТУ.