

На правах рукописи



Титов Сергей Борисович

**ГИБРИДНЫЕ АЛГОРИТМЫ  
АНАЛИЗА И ОБРАБОТКИ ДАННЫХ  
В ЗАДАЧАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ**

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

Автореферат  
диссертации на соискание ученой степени  
кандидата технических наук

Рязань 2011

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего профессионального образования «Рязанский государственный радиотехнический университет»

Научный руководитель: доктор технических наук, доцент  
ДЕМИДОВА Лилия Анатольевна

Официальные оппоненты: доктор технических наук, профессор  
СКВОРЦОВ Сергей Владимирович

кандидат технических наук, доцент  
МИНАЕВ Юрий Михайлович

Ведущая организация: Московский государственный открытый университет

Защита диссертации состоится 21 сентября 2011 г. в 12 часов на заседании диссертационного совета Д 212.211.01 в ФГБОУ ВПО «Рязанский государственный радиотехнический университет» по адресу: 390005, г. Рязань, ул. Гагарина, 59/1.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВПО «Рязанский государственный радиотехнический университет»

Автореферат разослан « \_\_\_\_ » \_\_\_\_\_ 2011 г.

Ученый секретарь  
диссертационного совета  
канд. техн. наук, доцент



В.Н. Пржегорлинский

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы исследования.** Проблема принятия решений присуща многим прикладным задачам, в том числе и задаче обработки информации в сфере городской жилой недвижимости (ГЖН). Эффективный подход к задаче поддержки принятия решений в современных условиях невозможен без привлечения новых информационных технологий, важной частью которых являются интеллектуальные средства обработки информации. В настоящее время одной из довольно актуальных проблем на рынке недвижимости является создание понятной и гибкой системы классификации объектов недвижимости. Российский рынок недвижимости, безусловно, нуждается в общепринятой и понятной потребителям классификации объектов. На сегодняшний день известно немало попыток создать подобную классификацию, однако ни одна из них не принята профессиональным сообществом в качестве стандарта. Наибольшие успехи в этом вопросе наблюдаются в отдельных сегментах рынка коммерческой недвижимости. Классификации рынка жилой недвижимости разрабатываются, как правило, внутри отдельно взятых компаний и плохо согласуются друг с другом.

Большую работу в данном направлении проделали такие крупные консалтинговые и девелоперские компании, как: Blackwood, МИЭЛЬ, Penny Lane Realty, Knight Frank, Калинка Риэлти, CB Richard Ellis, Swiss Realty Group, Российская гильдия риэлторов Ведис Групп, Capital Group, МИАН, Гута-Девелопмент, Дон-Строй, Уникор, БАРКЛИ и др.

Безусловно, создать предпосылки для формирования единого подхода при разработке методики классификации объектов ГЖН довольно сложно, особенно в условиях постоянного развития рынка. Отечественный рынок недвижимости сравнительно молод, но развивается стремительными темпами, причем порой даже в различных направлениях. Меняется подход к качеству строительства жилых зданий, внедряются перспективные разработки, меняются и требования покупателей к качеству и иным характеристикам такой продукции. Как следствие, и принадлежность объекта к тому или иному классу – величина ничуть не predetermined, она также должна пересматриваться во времени.

Существующие алгоритмы и методы классификации объектов ГЖН в своём большинстве базируются на учете фиксированного набора характеристик объекта, на основе анализа которого и делается вывод об отнесении объекта к тому или иному классу, однако при этом зачастую не учитывается динамическая составляющая рынка недвижимости. Более того, при решении задач классификации объектов ГЖН порой помимо объективных характеристик требуется учет и ряда субъективных характеристик, которым присуща некоторая неопределенность. Так, при сопоставлении неформализованных параметров объектов ГЖН, представляющих собой качественные характеристики объектов, целесообразно использовать экспертное оценивание, которое может оказаться особенно полезным в случаях выявления неполноты и неточности необходимых для анализа данных.

В то же время недостаточно просто дать оценки некоторому объекту ГЖН по выбранному набору характеристик: необходимо принять обоснованное и адекватное решение, позволяющее определить класс принадлежности объекта. В случае мониторинга рынка ГЖН классификация может быть выполнена посредством кластеризации множества объектов ГЖН для выбранного набора характеристик, например, на заданное или произвольное количество кластеров (классов).

Зачастую при решении задачи классификации объектов ГЖН возникает проблема последовательной поэтапной классификации, предполагающая проведение предварительной (базовой) классификации объектов по некоторому набору характеристик, последующую дополнительную классификацию выбранного класса объектов по какому-либо расширенному набору количественных и качественных характеристик и заключительное упорядочение объектов целевого подкласса по некоторому новому набору качественных характеристик, отражающих субъективные предпочтения потребителя.

Таким образом, существует необходимость в разработке новых альтернативных подходов к проблеме классификации объектов ГЖН. В настоящее время при решении широкого спектра прикладных задач всё большее внимание уделяется применению гибридных технологий, реализующих комплексное применение различных алгоритмов и методов искусственного интеллекта, разработанных в рамках теории искусственных иммунных систем (ИИС), теории нечетких множеств (ТНМ), теории генетических алгоритмов (ГА), теории искусственных нейронных сетей, теории мультимножеств (ТММ) и т.п.

Один из современных подходов, используемых в различных задачах обработки информации, основан на принципах функционирования ИИС. Способность ИИС к самоорганизации определяет возможность их применения в таких задачах, как, например, кластеризация объектов. Принципы самоорганизации были предметом исследования многих ученых и описаны в работах Дж. фон Неймана, Н. Винера, У.Р. Эшби и др.

Зачастую на практике задачи кластеризации предполагают принятие решений в условиях неопределенности. Применение ТНМ, основоположником которой является Л.А. Заде, позволяет строить формальные схемы решения задач, характеризующихся той или иной степенью неопределенности, обусловленной неполнотой, внутренней противоречивостью, неоднозначностью, субъективностью исходных данных, представляющих собой приближенные количественные или качественные оценки характеристик объектов. Анализ алгоритмов кластеризации, основанных на применении ТНМ, показывает, что довольно часто они не обеспечивают получение адекватных решений ввиду недостаточного обоснованного выбора их параметров, а поиск эффективных решений приводит к значительным временным затратам из-за необходимости выполнения многократных реализаций классических алгоритмов с целью выбора оптимальных параметров. В последние годы для решения оптимизационных задач, которые трудноразрешимы классическими методами, успешно применяются ГА – адаптивные методы поиска, реализующие 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. GOLдберга, В.В. Емельянова и др.

Совместное использование инструментария ТНМ и ТММ, основные положения которой изложены в работах А.Б. Петровского, обеспечивает принятие адекватных и обоснованных решений по классификации и упорядочению объектов ГЖН по некоторому набору субъективных характеристик в случае наличия несовпадающих и даже противоречивых исходных групповых экспертных оценок.

Использование комплексного подхода к решению задачи классификации объектов ГЖН с применением инструментария ИИС, ТНМ, ГА и ТММ позволит создать качественно новые программные средства, обеспечивающие для задач классификации объектов ГЖН повышение обоснованности и объективности принятия классификационных решений при приемлемых временных затратах.

**Цель диссертационной работы** состоит в повышении обоснованности принимаемых классификационных решений в сфере ГЖН посредством разработки эффективных гибридных алгоритмов кластеризации и упорядочения объектов ГЖН.

Для достижения поставленной цели необходимо решить следующие задачи.

1. Провести анализ существующих алгоритмов кластеризации и упорядочения объектов ГЖН, выявить перспективные направления их развития.

2. Исследовать возможность комплексного использования инструментария ИИС, ТНМ, ГА и ТММ при разработке алгоритмов кластеризации и упорядочения объектов ГЖН.

3. Разработать алгоритм кластеризации объектов ГЖН с использованием инструментария ИИС.

4. Разработать алгоритм кластеризации объектов ГЖН с использованием инструментария ТНМ и ГА.

5. Разработать алгоритмы упорядочения объектов ГЖН при групповом экспертном оценивании с использованием инструментария ТММ и ТНМ.

6. Разработать пакет прикладных программ (ППП) для классификации объектов ГЖН на основе разработанных алгоритмов кластеризации и упорядочения объектов.

**Методы исследования.** Теоретические исследования выполнены с использованием методов системного анализа, теории вероятностей, математической статистики, теории искусственных иммунных систем, теории нечетких множеств, теории мультимножеств, теории генетических алгоритмов, математического моделирования и объектно-ориентированного программирования.

**Научная новизна.** В рамках диссертационной работы были получены следующие результаты.

1. Разработан и исследован алгоритм кластеризации объектов ГЖН с использованием инструментария ИИС, реализующий выявление количества, структуры и топологии кластеров за счет механизма супрессии иммунной сети и обеспечивающий получение обоснованных результатов кластеризации множества объектов.

2. Разработан и исследован нечетко-возможностный алгоритм кластеризации объектов ГЖН на основе интервальных нечетких множеств второго типа (ИНМТ2) и ГА, позволяющий учесть свойства кластерной типичности и кластерной относительности и обеспечивающий получение обоснованных результатов кластеризации множества объектов, содержащего кластеры существенно разной плотности или существенно разного объема, с приемлемыми временными затратами в случае неопределенности выбора алгоритма кластеризации.

3. Разработан и исследован алгоритм упорядочения объектов ГЖН при групповом экспертном оценивании с использованием инструментария ТММ и ТНМ, позволяющий учесть все, в том числе противоречивые, оценки объектов ГЖН и обеспечивающий принятие обоснованных решений без использования дополнительных преобразований типа усреднения и смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных, с реализацией механизма интегрального согласования частных вариантов упорядочения объектов ГЖН.

4. Разработан и исследован алгоритм упорядочения объектов ГЖН с вычислением центроидов ИНМТ2 на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы, позволяющий дифференцировать объекты ГЖН, имеющие одинаковые центроиды нечетких множеств первого типа (НМТ1) на основе этих же экспертных оценок.

5. Разработана методика поэтапной классификации объектов ГЖН на произвольное количество кластеров с учетом произвольного количества объективных и субъективных характеристик с использованием разработанных алгоритмов кластеризации и упорядочения объектов ГЖН, реализующая комплексное решение задачи классификации объектов ГЖН.

**Практическая ценность работы.** Практическая ценность работы состоит в том, что разработанные алгоритмы кластеризации и упорядочения позволяют реализовать новый подход к задаче классификации объектов ГЖН и обеспечивают:

- «гибкость» применяемых алгоритмов анализа и обработки данных, способных учитывать динамику рынка и возможность изменения структуры классов объектов ГЖН;
- комплексный подход к задаче классификации объектов ГЖН, обеспечивающий последовательный учет ряда объективных и субъективных (количественных и качественных) характеристик оценивания;
- высокую обоснованность принятия решения в задачах классификации объектов

ГЖН в условиях неопределенности и неточности исходной информации, в том числе при несовпадающих (противоречивых) оценках экспертов;

– минимизацию временных и финансовых затрат, связанных как с необходимостью сбора и учета точных и полных исходных данных (что может быть принципиально невозможным), так и с необходимостью многократной реализации классических алгоритмов кластеризации с целью выбора соответствующих оптимальных параметров, обеспечивающих принятие адекватных решений.

В конечном итоге предлагаемый подход обеспечивает эффективное решение задачи классификации объектов ГЖН с привлечением субъективного «человеческого фактора».

**Достоверность** полученных в диссертационной работе результатов подтверждается:

- использованием понятий и выводов теории ИИС, ТНМ, теории ГА и ТММ;
- результатами математического моделирования предложенных алгоритмов на ПЭВМ;
- разработкой действующих программных средств, подтвержденных свидетельствами об официальной регистрации;
- апробацией предложенных алгоритмов и методики на конкретных примерах;
- наличием актов внедрения результатов диссертационной работы.

**На защиту выносятся:**

1. Алгоритм кластеризации объектов городской жилой недвижимости с использованием инструментария искусственных иммунных систем.

2. Нечетко-возможностный алгоритм кластеризации объектов городской жилой недвижимости на основе интервальных нечетких множеств второго типа и генетических алгоритмов.

3. Алгоритм упорядочения объектов городской жилой недвижимости при групповом экспертном оценивании с использованием инструментария теории множеств и теории нечетких множеств с реализацией механизма интегрального согласования частных вариантов упорядочения объектов городской жилой недвижимости.

4. Алгоритм упорядочения объектов городской жилой недвижимости с вычислением центроидов нечетких множеств первого типа и интервальных нечетких множеств второго типа на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы.

5. Методика поэтапной классификации объектов городской жилой недвижимости на произвольное количество кластеров с учетом произвольного количества объективных и субъективных характеристик с использованием разработанных алгоритмов кластеризации и упорядочения объектов городской жилой недвижимости.

6. ППП для классификации объектов городской жилой недвижимости на основе разработанных алгоритмов кластеризации и упорядочения.

**Внедрение результатов.** Результаты работы внедрены и используются в деятельности ООО «Фонд строительных инвестиций», а также в работе Рязанского регионального отделения Общероссийской общественной организации «Русское общество оценщиков» при анализе и классификации объектов ГЖН. Опытная эксплуатация ППП «Классификация недвижимости на основе гибридных технологий» показала высокие характеристики эффективности разработанного программного обеспечения при решении задач классификации совокупностей объектов ГЖН. Использование результатов диссертационной работы на практике подтверждено соответствующими актами о внедрении.

**Апробация работы.** Основные положения диссертации докладывались на следующих конференциях: 32-я всероссийская научно-техническая конференция «Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный государственный образовательный стандарт 3-го поколения» (апрель 2007 г., г. Рязань); 33-я всероссийская научно-техническая конференция «Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный государственный образователь-

ный стандарт 3-го поколения» (апрель 2008 г., г. Рязань); VIII Всероссийская научно-техническая конференция «Проблемы информатики в образовании, управлении, экономике и технике» (ноябрь 2008 г., г. Пенза); 34-я всероссийская научно-техническая конференция «Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный государственный образовательный стандарт 3-го поколения» (апрель 2009 г., г. Рязань); Всероссийская научная конференция «Современные исследовательские и образовательные технологии» (МАПР-09) (октябрь 2010 г., г. Таганрог); XVI международная открытая научная конференция «Современные проблемы информатизации в экономике и обеспечении безопасности», (ноябрь 2010 г. – январь 2011 г., г. Воронеж); VII международная научно-практическая конференция «Современные вопросы науки – XXI век» (март 2011 г., г. Тамбов).

**Публикации.** По теме диссертации опубликовано 15 печатных работ. В их числе 1 статья в рецензируемой печати, 4 статьи в межвузовских сборниках, 1 статья в научно-техническом журнале, 2 доклада на международных конференциях, 5 докладов на всероссийских конференциях, 1 свидетельство о регистрации программы для ЭВМ в ФГУ ФИПС – РОСПАТЕНТ, 1 свидетельство об официальной регистрации программы в Отраслевом фонде алгоритмов и программ (ОФАП).

**Структура и объем диссертации.** Диссертационная работа состоит из введения, пяти глав, списка литературы и трех приложений. Содержит 244 страницы (из них 232 страницы – основная часть, 12 страниц – приложения), 30 таблиц, 36 рисунков. Список литературы состоит из 187 наименований.

## СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обосновывается актуальность выбора темы диссертации, формулируются цель и задачи исследований, научная новизна и практическая ценность основных результатов диссертационной работы.

**В первой главе «Обзор и анализ подходов к решению задачи оценки и классификации объектов городской жилой недвижимости»** сформулирована задача разработки нового подхода к классификации объектов ГЖН, учитывающего реалии быстро развивающегося современного города. Показано, что задача классификации объектов ГЖН возникает достаточно часто и необходимость ее решения ощущается очень остро специалистами, работающими на рынке недвижимости, риэлторами, оценщиками, девелоперами и др. Проанализированы подходы к классификации объектов ГЖН, используемые на рынке недвижимости, выявлены их недостатки (громоздкость, избыточность и т.п.). Отмечено, что существующие на данный момент подходы к классификации ГЖН весьма далеки от совершенства. Сделан вывод о целесообразности разработки алгоритмов кластеризации на основе искусственной иммунной сети, а также на основе ИНМТ2 и ГА для решения задачи классификации объектов ГЖН.

Предложен новый подход к решению задачи классификации объектов ГЖН, основанный на комплексном применении инструментария ИИС, ТНМ, ГА и ТММ и предполагающий выполнение трех следующих этапов.

1. Базовая классификация (кластеризация) объектов ГЖН по заранее фиксированному набору наиболее важных (общепринятых) характеристик, отражающих качественные и количественные свойства объектов ГЖН, с использованием инструментария ИИС.

2. Дополнительная классификация (кластеризация) объектов ГЖН в пределах одного кластера (или нескольких уже сформированных кластеров) по каким-либо специфицированным наборам качественных и количественных характеристик, которые определяются уже для каждого конкретного случая в индивидуальном порядке, на основе алгоритма кластеризации, предполагающего комплексное использование алгоритмов нечеткой и возможностной кластеризации и выявление «лучшего» специфицированного подкластера (с точки зрения его близости к «идеальному» объекту недвижимости, гипотети-

чески обладающему самыми «лучшими» значениями по всем качественным и количественным характеристикам).

3. Упорядочение объектов ГЖН в пределах специфицированного подкластера объектов ГЖН с учетом ряда субъективных качественных характеристик, оценки по которым выставляются группой экспертов, с использованием инструментария ТММ и ТНМ.

Таким образом, можно говорить о наличии трех взаимосвязанных задач, связанных с классификацией объектов ГЖН: задачи базовой классификации объектов ГЖН, задачи дополнительной классификации объектов ГЖН, задачи упорядочения объектов ГЖН.

**Во второй главе «Разработка алгоритма кластеризации объектов городской жилой недвижимости с использованием механизма искусственных иммунных систем»** рассматривается проблема разработки алгоритма базовой классификации (кластеризации) объектов ГЖН по набору наиболее важных (общепринятых) характеристик, отражающих качественные и количественные свойства объектов недвижимости, представленных многоквартирными жилыми домами.

Для решения задачи базовой классификации (кластеризации) объектов ГЖН исследована возможность применения механизма ИИС для определения количества, структуры и топологии кластеров.

Пусть  $n_1$  – количество объектов ГЖН,  $q_1$  – количество характеристик.

Искусственная иммунная сеть представляет собой полносвязный граф, состоящий из множества узлов – антител сети и множества взвешенных ребер, устанавливающих связи между антителами. Задача антител иммунной сети заключается в распознавании антигенов, в качестве которых в контексте решаемой задачи базовой классификации (кластеризации) объектов выступают объекты ГЖН, представляемые в виде векторов  $x_i$  ( $i = \overline{1, n_1}$ ), состоящих из  $q_1$  элементов, соответствующих характеристикам оценивания. Все характеристики (в том числе и качественные) должны быть приведены к единому виду, то есть формализованы (например, с помощью экспертных оценок) и нормированы.

Алгоритм формирования искусственной иммунной сети может быть описан следующим образом:

$$immNET = (X, C, M, N_C, D, S, n, \xi, \sigma_d, \sigma_s), \quad (1)$$

где  $X$  – множество объектов  $x_i$  ( $i = \overline{1, n_1}$ ) размерностью  $q_1$ , определяющих популяцию антигенов;  $C$  – матрица памяти, содержащая все антитела сети ( $C \in R^{n_1 \times q_1}$ ), определяющие популяцию антител;  $M$  – матрица, состоящая из  $N$  антител памяти ( $M \subseteq C$ );  $N_C$  – общее количество клонов, создаваемых стимулируемыми антителами в каждом поколении;  $D$  – матрица значений аффинностей  $d_{ij}$  связей типа «антиген-антитело»  $Ag - Ab$ ;  $S$  – матрица значений аффинностей  $S_{ij}$  связей типа «антитело-антитело»  $Ab - Ab$ ;  $n$  – количество лучших антител, выбираемых из матрицы  $C$  для клонирования и мутации;  $\xi$  – процент улучшенных антител, отбираемых из популяции клонов для последующей обработки;  $\sigma_d$  – пороговый коэффициент подавления антитела в зависимости от значений аффинностей  $d_{ij}$  его связей типа «антиген-антитело»  $Ag - Ab$ ;  $\sigma_s$  – пороговый коэффициент супрессии сети.

В иммунных сетях различают два вида аффинности: аффинность связи «антиген-антитело»  $Ag - Ab$ , характеризующую степень расхождения антигена  $Ag$  и антитела  $Ab$ , и аффинность связи «антитело-антитело»  $Ab - Ab$ , характеризующую степень подобия двух антител  $Ab$ . Аффинитет определяется как число, принадлежащее отрезку  $[0,1]$ , и вычисляется как евклидово расстояние между двумя векторами, характеризующими два объекта, например антиген  $Ag$  и антитело  $Ab$ . Низкое значение аффинитета указывает на сильную близость двух объектов.



В предлагаемой иммунной сети антитела выполняют роль внутренних образов антигенов – реальных объектов ГЖН, используемых при построении иммунной сети.

Для решения проблемы определения структуры иммунной сети предложено использовать понятие «минимальное остовное дерево» (minimum spanning tree – MST) – остовное дерево с минимальным весом, где вес дерева равен сумме весов составляющих его ребер. MST является мощным механизмом поиска, реализующим локально адаптивную стратегию соединения антител иммунной сети, и может быть использовано для описания количества, структуры и топологии кластеров иммунной сети.

Анализ вычислительной сложности известных алгоритмов построения MST (алгоритма Борувки, алгоритма Крускала и алгоритма Прима) показал целесообразность использования для разбиения иммунной сети на кластеры на основе MST алгоритма Прима, характеризующегося более низкой вычислительной сложностью.

Зависимость качества классификации от размера иммунной сети является нелинейной, и начиная с некоторого момента, увеличение размера сети не дает существенного прироста качества классификации. Показано, что для оптимизации процесса классификации с использованием искусственной иммунной сети целесообразно управлять величиной степени сжатия иммунной сети – коэффициентом супрессии сети  $\sigma_s$ .

Для решения задачи поиска оптимального значения коэффициента супрессии  $\sigma_s$  в смысле достижения заданной степени сжатия иммунной сети предложено использовать наиболее простой вариант итеративного поиска:  $\sigma_s = \sigma_s \cdot V_{\sigma_s}$  для уменьшения коэффициента супрессии и  $\sigma_s = \sigma_s / V_{\sigma_s}$  для увеличения коэффициента супрессии, где  $V_{\sigma_s}$  – скорость изменения коэффициента супрессии;  $0 < V_{\sigma_s} < 1$  (например,  $V_{\sigma_s} = 0,95$ ). Начальное значение коэффициента супрессии полагается равным некоторому значению:  $\sigma_s = \sigma_{s_0}$ , а итерационные вычисления заканчиваются по достижении (с некоторой точностью  $\varepsilon$ ) заданного размера иммунной сети (количества антител)  $N_{max}$ .

Использование инструментария ИИС позволяет выполнить базовую классификацию (кластеризацию) объектов ГЖН, играющих роль антигенов, в результате которой выявляется такой минимальный набор антител, который отображает количество, структуру и топологию кластеров и однозначно определяет принадлежность любого из уже имеющихся объектов ГЖН к некоторому кластеру. Иммунная сеть может использоваться для выявления аномалий данных (например, как ошибочных данных, в которых допущены ошибки, подлежащие корректировке, так и принципиально новых данных), которые не могут быть классифицированы с помощью сформированной иммунной сети, поскольку их невозможно отнести ни к одному из кластеров, определяемых на основе MST.

Качество распознавания иммунной сетью множества антигенов характеризуется значением среднего аффинитета связей типа «антиген-антитело»  $Ag - Ab$  для всех антигенов, распознанных как «свои» антителами соответствующих кластеров: чем меньше данное значение, тем лучше антитела иммунной сети распознают антигены.

$$averageD = \left( \sum_{i=1}^{n_1} \min_{j=1, k} d_{ji} \right) / n_1, \quad (2)$$

где  $n_1$  – количество антигенов  $Ag_i$ ;  $k$  – количество антител  $Ab_j$ ;  $d_{ji}$  – аффинность связи типа «антиген-антитело»  $Ag - Ab$ , определяемая через евклидово расстояние между  $j$ -м антителом  $Ab_j$  и  $i$ -м антигеном  $Ag_i$  ( $j = 1, k$ ,  $i = 1, n_1$ ).

В общем случае условием окончания алгоритма формирования иммунной сети является достижение предопределенных количества смен поколений или значения среднего аффинитета вида (2). В модифицированной версии алгоритма формирования иммунной сети предложено дополнительно использовать условие достижения заданной степени сжатия иммунной сети. Исследование работы алгоритма формирования иммунной сети

на тестовых и реальных данных показало, что сжатие данных в большинстве случаев составляет порядка 20-70 %, а в отдельных случаях может достигать 90 %. Механизм сжатия сети позволяет ускорить работу алгоритма при распознавании новых объектов ГЖН и отнесении их к одному из существующих классов (кластеров). Разработанный алгоритм кластеризации был применен к реальной выборке объектов ГЖН, содержащей 89 объектов, представленных к продаже на сайте одного из рязанских агентств недвижимости. Кластеризация проводилась по основным ценообразующим характеристикам, определяющим класс объекта ГЖН: «местоположение»; «материал строительства»; «архитектурное решение»; «техническое оснащение»; «высота потолков»; «охрана/организация безопасности»; «инфраструктура»; «территория дома». Для данного примера степень «сжатия» сети составила 37 % при коэффициенте супрессии 0,2, поскольку множество объектов кластеризации содержит 89 объектов (антигенов), а размер сформированной иммунной сети равен 33. Полученные результаты кластеризации (табл. 1, рис. 1) вполне объяснимы и логичны, что свидетельствует о корректности реализации алгоритма кластеризации с использованием механизма ИИС.



Рисунок 1 – Результаты кластеризации (1-й этап)

Таблица 1 – Результаты классификации (кластеризации) с использованием механизма ИИС

Объекты ГЖН (антигены)			
1-й класс	2-й класс	3-й класс	4-й класс
30, 34	16, 29, 51, 77	2, 3, 5, 6, 7, 8, 9, 10, 13, 14, 15, 17, 18, 20, 21, 22, 23, 25, 27, 28, 31, 33, 36, 38, 39, 40, 44, 46, 47, 48, 49, 50, 52, 53, 55, 56, 57, 61, 67, 73, 74, 75, 78, 79, 80, 81, 82, 85, 86, 87, 88, 89	1, 4, 11, 12, 19, 24, 26, 32, 35, 37, 41, 42, 43, 45, 54, 58, 59, 60, 62, 63, 64, 65, 66, 68, 69, 70, 71, 72, 76, 83, 84
Для распознавания антигенов класса сформировано 2 антитела	Для распознавания антигенов класса сформировано 3 антитела	Для распознавания антигенов класса сформировано 20 антител	Для распознавания антигенов класса сформировано 8 антител

Отмечено, что использование предложенного алгоритма кластеризации позволяет повысить обоснованность принимаемых решений в ходе анализа и обработки информации в сфере ГЖН по сравнению с существующими на текущий момент на региональном рынке недвижимости зачастую интуитивными вариантами классификации за счет использования нового подхода к задаче классификации объектов ГЖН, основанного на применении адекватного математического аппарата.

**В третьей главе «Разработка алгоритма кластеризации объектов городской жилой недвижимости в условиях неопределенности выбора алгоритма кластеризации»** рассматривается проблема разработки алгоритма кластеризации множества объектов ГЖН, содержащего кластеры существенно разного объема или существенно разной плотности, в условиях неопределенности выбора алгоритмов нечеткой и возможностной кластеризации на основе дополнительного набора специфицированных характеристик.

Пусть  $n_2$  – количество объектов ГЖН,  $q_2$  – количество характеристик.

Для выполнения кластеризации объектов ГЖН могут быть использованы алгоритм нечетких  $c$ -средних на основе НМТ1 и алгоритм возможных  $c$ -средних на основе НМТ1, реализующие соответственно учет свойства кластерной относительности объектов и свойства кластерной типичности объектов и обеспечивающие получение адекватных результатов кластеризации множеств объектов, содержащих кластеры подобной плотности и подобного объема. Алгоритм нечетких  $c$ -средних на основе НМТ1 реализует минимизацию целевой функции вида:

$$J(U, V) = \sum_{j=1}^c \sum_{i=1}^{n_2} (u_j(x_i))^m \cdot d_{ji}^2 \left( d_{ji}^2 = \sum_{l=1}^{q_2} (x_i^l - v_j^l)^2 \right), \quad (3)$$

где  $U = [u_j(x_i)]$  – нечеткое  $c$ -разбиение множества объектов  $\{x_i\}$  на основе функций принадлежности  $u_j(x_i)$ ;  $V = (v_1, \dots, v_c)$  – центры кластеров;  $d_{ji}$  – расстояние между центром кластера  $v_j$  и объектом  $x_i$ :  $d_{ji} = \|x_i - v_j\|$ ;  $m$  – фаззификатор ( $m \in R$ ,  $m > 1$ );  $c$  – количество кластеров;  $n_2$  – количество объектов кластеризации;  $q_2$  – количество характеристик.

Алгоритм нечетких  $c$ -средних на основе НМТ1 предполагает инициализацию начального нечеткого  $c$ -разбиения  $U = [u_j(x_i)]$  и итерационное уточнение координат центров кластеров:

$$v_j = \frac{\sum_{i=1}^{n_2} (u_j(x_i))^m \cdot x_i}{\sum_{i=1}^{n_2} (u_j(x_i))^m} \quad (4)$$

и значений функций принадлежности (ФП):

$$u_j(x_i) = 1 / \left( \sum_{k=1}^c (d_{ji} / d_{ki})^{2/(m-1)} \right) \left( \sum_{j=1}^c u_j(x_i) = 1, \forall i = \overline{1, n_2} \right) \quad (5)$$

до тех пор, пока не будет выполнено заданное число итераций  $s$  или не будет достигнута заданная точность:  $|J(U, V) - J'(U, V)| \leq \varepsilon$ , где  $J(U, V)$ ,  $J'(U, V)$  – значения целевой функции на двух последовательных итерациях.

Алгоритм возможных  $c$ -средних на основе НМТ1 инициализируется с помощью нескольких итераций алгоритма нечетких  $c$ -средних на основе НМТ1 и реализует минимизацию целевой функции вида:

$$J(W, V) = \sum_{j=1}^c \sum_{i=1}^{n_2} (w_j(x_i))^m \cdot d_{ji}^2 + \sum_{j=1}^{\bar{n}} \eta_j^2 \cdot \sum_{i=1}^{n_2} (1 - w_j(x_i))^m, \quad (6)$$

где  $W = [w_j(x_i)]$  – возможностное  $c$ -разбиение множества объектов  $\{x_i\}$  на основе функций типичности  $w_j(x_i)$ ;  $V = (v_1, \dots, v_c)$  – центры кластеров;  $d_{ji}$  – расстояние между центром кластера  $v_j$  и объектом  $x_i$ ;  $m$  – фаззификатор;  $\eta_j$  ( $j = 1, c$ ) – «ширина зонь», определяющая расстояние, на котором значение функции типичности объекта  $j$ -му кластеру равно 0,5;  $c$  – количество кластеров;  $n_2$  – количество объектов кластеризации.

Функции типичности (ФТ) могут быть вычислены как

$$w_j(x_i) = 1 / \left( 1 + (d_{ji} / \eta_j)^{2/(m-1)} \right) \left( \sum_{i=1}^{n_2} w_j(x_i) = 1, \forall j = \overline{1, c} \right), \quad (7)$$

а координаты центров кластеров находятся в соответствии с формулой (4) при соответствующей замене ФП на ФТ.

Если кластеры в множестве объектов имеют существенно разную плотность или существенно разный объем, то работа алгоритмов кластеризации на основе НМТ1 существенно зависит от выбора параметров алгоритмов кластеризации. Кроме того, зачастую имеет место неопределенность в выборе того или иного алгоритма кластеризации, а значит, неопределенность в задании ФП и ФТ. Таким образом, существует необходимость в разработке алгоритма управления неопределенностью значений степеней принадлежности и

степеней типичности объектов кластерам, который бы одновременно учитывал свойство кластерной относительности для алгоритма нечетких  $c$ -средних на основе НМТ1 и свойство кластерной типичности для алгоритма возможностных  $c$ -средних на основе ИНМТ1.

Расширение множества объектов на ИНМТ2 позволяет управлять неопределенностью, которая возникает, когда множество объектов содержит кластеры существенно разной плотности или существенно разного объема. В связи с этим предлагается реализовать расширение множества объектов кластеризации на ИНМТ2, обеспечив управление неопределенностью посредством одновременного использования алгоритма нечетких  $c$ -средних на основе НМТ1 и алгоритма возможностных  $c$ -средних на основе ИНМТ1, что позволит существенно улучшить результаты кластеризации. При определении «отпечатка неопределенности» ИНМТ2 «верхняя» и «нижняя» ФП для объекта могут быть представлены с помощью ФП и ФТ вида (5) и (7) соответственно следующим образом:

$$\bar{u}_j(x_i) = \begin{cases} u_j(x_i), & \text{если } u_j(x_i) > w_j(x_i) \\ w_j(x_i), & \text{если } u_j(x_i) \leq w_j(x_i) \end{cases}, \quad \underline{u}_j(x_i) = \begin{cases} u_j(x_i), & \text{если } u_j(x_i) \leq w_j(x_i) \\ w_j(x_i), & \text{если } u_j(x_i) > w_j(x_i) \end{cases}, \quad (8)$$

где  $u_j(x_i)$  и  $w_j(x_i)$  определяются в соответствии с формулами (5) и (7).

В качестве целевых функций разрабатываемого алгоритма кластеризации на основе ИНМТ2 следует рассматривать целевые функции алгоритма нечетких  $c$ -средних на основе НМТ1 и алгоритма возможностных  $c$ -средних на основе ИНМТ1 вида (3) и (6) соответственно, в которых фаззификатор  $m$  принимает одно и то же значение.

В общем случае в качестве показателя качества кластеризации, который должны быть минимизированы, целесообразно использовать общий гиперобъем:

$$H = \sum_{j=1}^c (det(R_j))^{1/2}, \quad (R_j = \frac{1}{n_j} \cdot \sum_{i=1}^{n_j} (x_i - v_j) \cdot (x_i - v_j)^T), \quad (9)$$

где  $v_j$  – вектор координат центра  $j$ -го кластера;  $x_i$  – вектор координат  $i$ -го объекта;  $n_2$  – 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  $R_j$ .

Управление неопределенностью выбора фаззификатора  $m$  и «ширины зоны»  $\eta_j$  ( $j = \overline{1, c}$ ) в нечетко-возможностном алгоритме кластеризации на основе ИНМТ2 осуществляется с помощью процедур вычисления центров кластеров и дефаззификации (получения четкого разбиения). При оценке координат центра кластера используется итерационный алгоритм Карника – Мендела.

Для поиска оптимальной комбинации значения фаззификатора  $m$  и значений «ширины зоны»  $\eta_j$  ( $j = \overline{1, c}$ ) в нечетко-возможностном алгоритме кластеризации на основе ИНМТ2 предлагается использовать ГА. При этом хромосома задается в виде  $s = (m, \eta_1, \dots, \eta_c)$ , где  $m \in (1, m_{max})$ ;  $\eta_j \in [\eta_j^{min}, \eta_j^{max}]$ ;  $\eta_j^{min} \geq 0$ ,  $\eta_j^{max} \geq 0$ ,  $\eta_j^{min} \leq \eta_j^{max}$  ( $j = \overline{1, c}$ );  $m_{max}$  и  $\eta_{max}$  – некоторые действительные числа.

Сравнительный анализ результатов кластеризации, полученных при решении тестовых задач кластеризации множеств объектов, содержащих кластеры существенно разной плотности или существенно разного объема, с использованием алгоритма нечетких  $c$ -средних на основе НМТ1, алгоритма возможностных  $c$ -средних на основе ИНМТ1 и разработанного нечетко-возможностного алгоритма кластеризации на основе ИНМТ2 и ГА, показал, что применение разработанного нечетко-возможностного алгоритма кластеризации на основе ИНМТ2 и ГА позволяет уменьшить количество ошибочно классифицированных объектов на 4-10 %.

Разработанный нечетко-возможностный алгоритм кластеризации на основе ИНМТ2 и

ГА был применен для выполнения дополнительной классификации (кластеризации) объектов ГЖН, отнесенных к 4-му классу (кластеру), полученному при выполнении базовой классификации (кластеризации) реальной выборки объектов ГЖН (глава 2). На втором этапе классификации все дополнительные характеристики, влияющие на предпочтения участников рынка («общая площадь квартиры»; «количество комнат»; «расположение квартиры»; «этаж расположения квартиры»; «физическое состояние квартиры»; «удачность планировки»; «обеспеченность коммуникациями и парковочными местами»), были формализованы и приведены к сопоставимому виду с помощью экспертных оценок.

Лучшие результаты кластеризации (количество кластеров и соответственно разбиение объектов на кластеры), определенные по условию достижения минимального значения показателя качества кластеризации, имеют структуру, представленную на рисунке 2, и позволяют говорить о корректности реализации нечетко-возможностного алгоритма кластеризации на основе ИНМТ2 и ГА при выполнении дополнительной классификации (кластеризации) объектов ГЖН. Все подкластеры содержат сходные по заявленным характеристикам жилые помещения. Ошибка кластеризации составила 1 объект (№13), который был ошибочно отнесен к подкластеру №3. Данный объект представляет собой жилое помещение, расположенное в п.Солотча, таким образом, можно сказать, что этот объект держится особняком среди прочих объектов (является шумовым объектом) и его ошибочное отнесение к подкластеру №4 вызвано именно данным фактом.

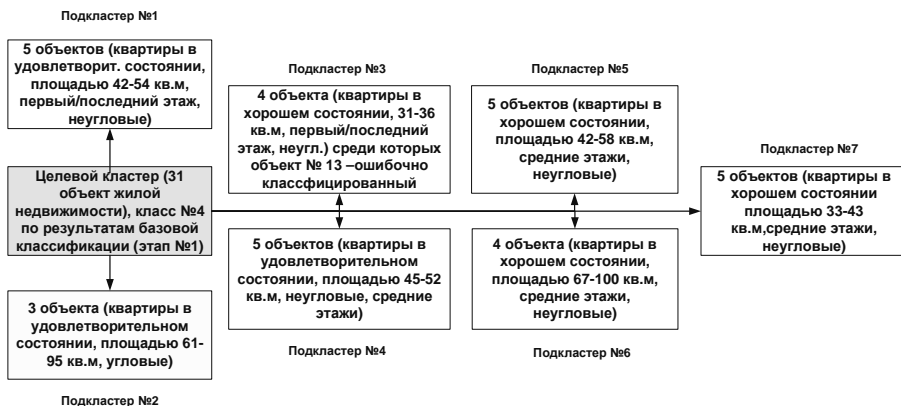


Рисунок 2 – Результаты кластеризации (2-й этап)

Предлагаемый нечетко-возможностный алгоритм кластеризации на основе ИНМТ2 и ГА в условиях неопределенности выбора алгоритма кластеризации реализует управление неопределенностью, возникающей при анализе множества объектов, содержащего кластеры существенно разной плотности или объема, а применение ГА позволяет при приемлемых временных затратах найти оптимальную комбинацию значений параметров нечетко-возможностного алгоритма кластеризации: фаззификатора  $m$  и «ширины зоны»  $\eta_j$  ( $j = \overline{1, c}$ ), обеспечивающую получение адекватных результатов кластеризации.

**В четвертой главе «Разработка алгоритмов упорядочения объектов с использованием экспертного оценивания»** рассматривается проблема разработки алгоритмов упорядочения множества объектов ГЖН на основе их свойств, выраженных значениями оценок объектов ГЖН по некоторому набору характеристик (отражающих субъективные предпочтения потребителя), при групповом экспертном оценивании.

Пусть  $n_3, q_3, M$  – количество объектов ГЖН, характеристик и экспертов соответственно. Для представления объектов ГЖН, оцениваемых группой экспертов по некоторому

набору характеристик, была выбрана математическая модель – «мультимножество» (ММ), использование которой позволяет учесть все комбинации значений оценок по характеристикам, а также количество значений оценок по каждой из этих характеристик (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 характеристикам для описания объектов ГЖН.

Пусть по каждой  $l$ -й характеристике ( $l = \overline{1, q_3}$ ) экспертами было дано  $z_l$  различных значений оценок  $p_l^{y_l}$  ( $y_l = \overline{1, z_l}$ ). Предположим, что все значения оценок  $p_l^{y_l}$  ( $l = \overline{1, q_3}$ ,  $y_l = \overline{1, z_l}$ ) независимы и упорядочены от худшего значения к лучшему:  $p_l^1 < \dots < p_l^{z_l}$ .

Пусть, кроме того, количество экспертов, давших объекту ГЖН  $x_i$  значение оценки  $p_l^{y_l}$ , равно  $k_{x_i}(p_l^{y_l})$  и  $\sum_{y_l=1}^{z_l} k_{x_i}(p_l^{y_l}) = M$  ( $i = \overline{1, n_3}$ ,  $l = \overline{1, q_3}$ ). Для каждого объекта  $x_i$  ( $i = \overline{1, n_3}$ ) на множестве характеристик  $P = \{P_1, P_2, \dots, P_{q_3}\}$  формируется ММ вида:

$$s_i = \{k_{x_i}(p_1^1) \bullet p_1^1, \dots, k_{x_i}(p_1^{z_1}) \bullet p_1^{z_1}, \dots, k_{x_i}(p_{q_3}^1) \bullet p_{q_3}^1, \dots, k_{x_i}(p_{q_3}^{z_{q_3}}) \bullet p_{q_3}^{z_{q_3}}\}, \quad (10)$$

где кратность  $k_{x_i}(p_l^{y_l})$  ММ характеризует количество экспертов, давших объекту  $x_i$ , которому соответствует ММ  $s_i$ , оценку  $p_l^{y_l}$ .

Гипотетическому «идеальному» объекту ГЖН соответствует ММ вида

$$s_{max} = \{M \bullet p_1^1, 0, \dots, 0, M \bullet p_2^1, 0, \dots, 0, M \bullet p_{q_3}^1, 0, \dots, 0\}. \quad (11)$$

В таком случае задача упорядочивания объектов ГЖН по некоторому набору характеристик сводится к упорядочению ММ по близости к «идеальному» объекту ГЖН в метрическом пространстве ММ  $(S, d)$  с метрикой Хемминга:

$$d = m(ADB) = \sum_{l=1}^{q_3} \chi_l \cdot \sum_{y_l=1}^{z_l} |k_A(p_l^{y_l}) - k_B(p_l^{y_l})|, \quad (12)$$

где  $A, B$  – ММ,  $\chi_l > 0$  – коэффициент относительной важности  $l$ -й характеристики ( $l = \overline{1, q_3}$ ).

При этом сначала выполняется сравнение взвешенных сумм значений первых оценок объектов по характеристикам:  $H_{s_h}^1 = \sum_{l=1}^{q_3} \chi_l \cdot k_{x_h}(p_l^1)$ . Лучшим будет объект ГЖН  $x_h$ , для которого сумма  $H_{s_h}^1$  будет наибольшей. Для упорядочения объектов ГЖН внутри группы эквивалентных объектов ГЖН  $x_h$  ( $h = \overline{1, r}$ ), имеющих равные суммы  $H_{s_h}^1$ , вычисляются взвешенные суммы значений вторых оценок объектов по характеристикам:

$H_{s_{h1}}^2 = \sum_{l=1}^{q_3} \chi_l \cdot k_{x_{h1}}(p_l^2)$ . Процесс вычисления взвешенных сумм продолжается до полного упорядочения всех объектов ГЖН.

При групповом экспертном оценивании объектов ГЖН по некоторому набору характеристик каждый эксперт может производить оценивание посредством: 1) прямого метода оценивания объектов ГЖН с применением какой-либо балльной шкалы; 2) метода парных сравнений объектов ГЖН по шкале Саати; 3) метода нечетких экспертных оценок объектов ГЖН, изначально заданных с помощью ФП (треугольных либо гауссовских).

В случае одновременной реализации более одного метода оценивания объектов ГЖН будет получено соответствующее количество частных вариантов упорядочения объектов ГЖН, представленных ММ, после обобщения которых может быть получен интеграль-

ный (согласующий) вариант упорядочения объектов ГЖН. Для получения интегрального варианта упорядочения объектов ГЖН предложено использовать метод медиан рангов (при отсутствии существенных противоречий среди согласуемых вариантов) либо метод согласования кластеризованных ранжировок (при наличии противоречий).

Предложенный алгоритм упорядочения был применен для реальной выборки объектов ГЖН, отнесенных к 5-му целевому подкластеру при классификации с применением алгоритма кластеризации в условиях неопределенности (глава 3). Характеристики объектов ГЖН (5 предложений о продаже ГЖН на рязанском рынке), попавших в целевой подкластер и подлежащих упорядочению, приведены в таблице 2. Все субъективные характеристики, влияющие на предпочтения вполне конкретного потребителя (а именно: репутация, имидж объекта; удобство добираться до места работы; удобство добираться до школы; оценка экологии в районе местоположения объекта), были формализованы и приведены к сопоставимому виду с помощью экспертных оценок. При упорядочении объектов ГЖН из таблицы 2 был учтен тот факт, что потребитель имеет личный автотранспорт, его место работы находится в центре города на ул. Семинарская; у него есть дети, которые ходят в начальные классы средней школы. Расстояния от данных объектов ГЖН до места работы учитывалось по данным сервиса Яндекс, также учитывалось окружение анализируемых объектов ГЖН, в частности близкое расположение школ.

Таблица 2 – Характеристики объектов ГЖН, попавших в целевой подкластер

№	Район	Улица	Комн.	Материал	Этаж	Площ. ОЖК	Сан. уз.	Цена, руб.
1	Нед	Сельских строителей	2-улу	блоч.	9/10	50/30/9	раз.	1 850 000
2	Цен	Лермонтова	2-брезж	пан.	2/5	45/29/6	сов.	1 600 000
3	Д-П	Новоселов	2-улу	пан.	3/9	52/33/9	раз.	2 200 000
4	Д-П	Новоселов	2-улу	пан.	8/9	54/31/8	раз.	1 980 000
5	Д-П	Зубковой	2-улу	пан.	6/10	52/28/8	раз.	1 980 000

Таблица 3 – Результаты упорядочения объекта от лучшего к худшему (3-й этап)

Ранги	1	2	3	4	5
1-й вариант упорядочения	№2	№3	№4	№5	№1
2-й вариант упорядочения	№2	№4	№3	№5	№1
3-й вариант упорядочения	№2	№3	№4	№1	№5
<b>Объекты</b>	<b>№1</b>	<b>№2</b>	<b>№3</b>	<b>№4</b>	<b>№5</b>
Медианы рангов	5	1	2	3	4
<b>Ранги</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>Интегральный результат упорядочения объектов по методу медиан рангов</b>	<b>№2</b>	<b>№3</b>	<b>№4</b>	<b>№5</b>	<b>№1</b>

Экспертные оценки выставлялись 3-мя независимыми экспертами с помощью метода прямого оценивания и метода парных сравнений с использованием шкалы Саати. В результате представления объектов ГЖН в виде мультимножеств и проведения процедуры упорядочения по близости к «идеальному» объекту ГЖН при использовании прямого метода оценивания и метода парных сравнений были получены соответственно следующие результаты упорядочения: №2, №3, №4, №5, №1 и №2; №4; №3; №5; №1. Кроме того, был реализован метод комплексного прямого экспертного упорядочения объектов ГЖН, в результате применения которого были получены следующие результаты упорядочения: №2, №3, №4, №1, №5. Ввиду отсутствия существенных противоречий среди трех согласуемых вариантов для реализации механизма интегрального варианта упорядочения объектов ГЖН был использован метод медиан рангов (табл. 3).

Приведенный пример подтвердил обоснованность и объективность принимаемых решений по упорядочению объектов ГЖН, представленных мультимножествами, на основе алгоритма упорядочения объектов ГЖН при групповом экспертном оценивании с использованием инструментария ТММ и ТНМ с реализацией механизма интегрального согласования частных вариантов упорядочения объектов ГЖН. При этом в процессе принятия решения учитываются все, в том числе и противоречивые, оценки экспертов.

В качестве альтернативного варианта оценивания объектов ГЖН группой экспертов при

выполнении их упорядочения по некоторому набору характеристик предлагается использовать центроиды НМТ1 и ИНМТ2, что позволит дифференцировать объекты ГЖН, имеющие одинаковые центроиды НМТ1, вычисляемые на основе групповых экспертных оценок по характеристикам, но различающиеся центроиды ИНМТ2. Рекомендовано для частичного согласования групповых экспертных оценок использовать нечеткий метод Дельфы, являющийся расширением классического четкого метода Дельфы. При этом для задания НМТ1 целесообразно использовать треугольные или гауссовские ФП:

Таблица 4 – Функции принадлежности, используемые для задания нечетких оценок

Треугольная	Гауссовская симметричная
$u(x) = \begin{cases} 0, & x \leq a_1 \\ (x - a_1)/(a_M - a_1), & a_1 \leq x \leq a_M \\ (a_2 - x)/(a_2 - a_M), & a_M \leq x \leq a_2 \\ 0, & a_2 \leq x \end{cases} \quad (a_1 \leq a_M \leq a_2)$	$u(x) = e^{-\frac{(x-b)^2}{c^2}}$

Использование ФП в соответствии с вышеприведенными формулами (табл. 4) позволяет упростить процедуру согласования групповых экспертных оценок с применением нечеткого метода Дельфы, поскольку туры согласования сопровождаются переопределением только трёх или двух параметров соответственно для каждого эксперта.

Так, при групповом экспертном оценивании объектов ГЖН по некоторому набору характеристик с помощью треугольных нечетких чисел (ТНЧ), заданных с помощью треугольных ФП и представляющих собой кортежи из 3 элементов (минимального, наиболее правдоподобного и максимального значений экспертных оценок)  $A_r = (a_1^r, a_M^r, a_2^r)$  ( $r = \overline{1, M}$ ), в ходе нескольких туров согласования с использованием нечеткого метода Дельфы определяются «частично» согласованные ТНЧ экспертов и среднее ТНЧ  $A_{average} = (m_1, m_M, m_2)$ , в результате дефазификации которого по методу центра тяжести определяются интегральные оценки – центроиды НМТ1. В случае если объекты ГЖН имеют одинаковые оценки на основе центроидов НМТ1, предлагается использовать центроиды ИНМТ2, вычисляемые на основе «отпечатков неопределенности» с использованием итерационного алгоритма Карника – Менделя. Аналогичным образом выполняется оценивание объектов ГЖН на основе гауссовских ФП (табл. 4) с последующим согласованием кортежей  $(b^r, c^r)$  ( $r = \overline{1, M}$ ) с использованием нечеткого метода Дельфы и вычислением центроидов НМТ1 и ИНМТ2. Кроме того, предложено, ввиду достаточно простого способа задания и «большей» гладкости гауссовских ФП, аппроксимировать получаемые верхнюю и нижнюю границы «отпечатка неопределенности» (рис. 3) более простыми математическими функциями, что позволит упростить процедуру вычисления обобщенного центроида ИНМТ2 мнений экспертов. При этом нижняя и верхняя границы «отпечатка неопределенности» аппроксимируются соответственно с использованием гауссовской симметричной ФП (табл. 4) и гауссовской двухсторонней ФП (13).

$$u(x) = \begin{cases} e^{-\frac{(x-b_1)^2}{c_1^2}}, & x < b_1, \\ 1, & b_1 < x < b_2 \\ e^{-\frac{(x-b_2)^2}{c_2^2}}, & x > b_2, \end{cases} \quad (b_1 < b_2), \quad (13)$$

где  $b_1, b_2$  – координаты центров максимума;  
 $c_1, c_2$  – коэффициенты концентрации.

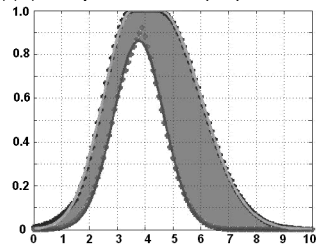


Рисунок 3 – «Отпечаток неопределенности»



Подбор оптимальных параметров аппроксимирующих функций реализован с привлечением ГА. При этом хромосома задается как  $s = (a, b, c, b_1, c_1, b_2, c_2)$ , где  $a$  – амплитуда «нижней» огибающей «отпечатка неопределенности», заданной с помощью ФП вида (14). Целевая функция ГА, которая должна быть минимизирована, задается как

$$F(a, b, c, b_1, c_1, b_2, c_2) = \sum_{l=1}^k (f_1(x_l) - u_1(x_l))^2 + \sum_{l=1}^k (f_2(x_l) - u_2(x_l))^2, \quad (14)$$

где  $f_1(x_l)$  и  $f_2(x_l)$  – соответственно значения «отпечатка неопределенности» по «нижней» и «верхней» границам, а  $u_1(x_l)$  и  $u_2(x_l)$  – соответственно значения гауссовских ФП (табл. 4) (с учетом амплитуды  $a$ ) и (13) в точках дискретизации  $x_l$  ( $l = \overline{1, k}$ ).

Практика применения предлагаемого подхода к аппроксимации «отпечатков неопределенности» при решении задач группового экспертного оценивания объектов ГЖН показала хорошую точность аппроксимации при малых вычислительных затратах.

**В пятой главе «Программная реализация гибридных алгоритмов анализа и обработки данных в сфере городской жилой недвижимости»** разработаны программные средства, предназначенные для выполнения комплексной поэтапной классификации объектов ГЖН по ряду характеристик оценивания на основе разработанных алгоритмов кластеризации и упорядочения объектов. ППП может быть рекомендован к использованию специалистами, работа которых связана с анализом и классификацией различных групп объектов, например объектов недвижимости и т.п. Комплексы программ, входящие в ППП, защищены авторским правом: получено свидетельство о регистрации программы для ЭВМ в ФГУ ФИПС (РОСПАТЕНТ) и свидетельство об официальной регистрации программ для ЭВМ в Отраслевом фонде алгоритмов и программ.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Проведенные исследования позволяют сформулировать основные результаты.

1. Проведен системный анализ традиционного подхода к решению задачи классификации объектов ГЖН, выявлены основные характеристики, по которым возможно осуществить классификацию объектов ГЖН. Отмечено, что существующие на текущий момент подходы к классификации ГЖН далеки от совершенства. Проведенный анализ теоретических аспектов классификации, кластеризации и упорядочения объектов ГЖН на основе значений их характеристик показал, что в случае группового экспертного оценивания объектов ГЖН при решении задачи упорядочения в качестве математической модели для представления объектов ГЖН может быть использовано мультимножество, позволяющее учесть все, в том числе и несовпадающие, значения характеристик объектов ГЖН. Сделан вывод о целесообразности использования алгоритмов кластеризации на основе ИИС, а также на основе ИНМТ2 и ГА для решения задачи классификации объектов ГЖН.

2. Разработан и исследован алгоритм кластеризации объектов ГЖН с использованием инструментария искусственных иммунных систем, реализующий выявление количества, структуры и топологии кластеров за счет механизма супрессии иммунной сети и обеспечивающий получение обоснованных результатов кластеризации множества объектов.

3. Разработан и исследован нечетко-возможностный алгоритм кластеризации объектов ГЖН на основе интервальных нечетких множеств второго типа и генетического алгоритма, позволяющий учесть свойства кластерной типичности и кластерной относительности и обеспечивающий получение обоснованных результатов кластеризации множества объектов, содержащего кластеры существенно разной плотности или существенно разного объема, с приемлемыми временными затратами в случае неопределенности выбора алгоритма кластеризации.

4. Разработан и исследован алгоритм упорядочения объектов ГЖН при групповом

экспертном оценивании с использованием инструментария теории множеств и теории нечетких множеств, позволяющий учесть все, в том числе противоречивые, оценки объектов ГЖН и обеспечивающий принятие обоснованных решений без использования дополнительных преобразований типа усреднения и смешивания, которые могут привести к необоснованным и необратимым искажениям исходных данных, с реализацией механизма интегрального согласования частных вариантов упорядочения объектов ГЖН.

5. Разработан и исследован алгоритм упорядочения объектов ГЖН с вычислением центроидов ИНМТ2 на основе экспертных оценок, согласованных с использованием нечеткого метода Дельфы, позволяющий дифференцировать объекты ГЖН, имеющие одинаковые центроиды НМТ1 на основе этих же экспертных оценок.

6. Разработана методика поэтапной классификации объектов ГЖН на произвольное количество кластеров с учетом произвольного количества объективных и субъективных характеристик с использованием разработанных алгоритмов кластеризации и упорядочения объектов ГЖН, реализующая комплексное решение задачи классификации объектов.

7. Разработан пакет прикладных программ для классификации объектов ГЖН на основе разработанных алгоритмов кластеризации и упорядочения.

## ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Демидова Л. А., Титов С.Б. Подход к проблеме нечеткой кластеризации в условиях неопределенности выбора целевой функции // Вестник РГРТУ. – Рязань, 2009. – №3 (выпуск 29). – С. 54-60 (в издании, рекомендованном ВАК РФ).

2. Демидова Л.А., Титов С.Б. Исследование проблемы согласования результатов при определении рыночной стоимости объекта оценки. Применение аппарата теории множеств в рамках оценочной деятельности // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. – М.: Горячая линия – Телеком, 2007. – С. 40-44.

3. Демидова Л.А., Титов С.Б. Классификация объектов городской жилой недвижимости с использованием алгоритма нечетких  $c$ -средних на основе нечетких множеств второго типа // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. – М.: Горячая линия – Телеком, 2008. – С. 129-135.

4. Демидова Л.А., Титов С.Б. Комплекс программ формирования обобщенного мнения экспертов на основе центроидов интервальных непрерывных нечетких множеств второго типа / Свидетельство об отраслевой регистрации разработки в ОФАП № 10623 от 12.07.2008 г.

5. Демидова Л.А., Титов С.Б. Комплексное применение инструментария искусственных иммунных систем, методов нечеткой кластеризации и теории множеств при решении задачи классификации объектов городской жилой недвижимости // Информационные технологии моделирования и управления. Научно-технический журнал. – Воронеж: «Научная книга», 2010. – №6(65). – С. 718-725.

6. Демидова Л.А., Титов С.Б. Классификация недвижимости на основе гибридных технологий (RealEstateClassificationOnHybridTechnologies) / Свидетельство о регистрации программы для ЭВМ в ФГУ ФИПС (РОСПАТЕНТ) №2011613199 от 22 апреля 2011 г.

7. Титов С.Б. Применение нечеткого метода Дельфы для оценки влияния экологических факторов на стоимость недвижимости // Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный образовательный стандарт 3-го поколения: материалы 32-й Всероссийской научно-технической конференции. – Рязань: Рязанское высшее военное командное училище связи, 2007. – С. 249-251.

8. Титов С.Б. Применение генетических алгоритмов при гауссовской аппроксимации функций принадлежности экспертных оценок // Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный образовательный стандарт 3-го поколения: материалы 33-й Всероссийской научно-технической конференции. – Рязань: Рязанское высшее военное командное училище связи, 2008. – С. 239-241.

9. Титов С.Б. Применение теории нечетких множеств второго типа к решению задачи классификации жилой недвижимости // Проблемы информатики в образовании, управлении, экономике и технике: материалы VII Всероссийской научно-технической конференции. – Пенза: Пензенский государственный педагогический университет им. В.Г.Белинского, 2008. – С.263-265.

10. Титов С.Б. Применение математического аппарата упорядочения множеств для ре-

шения задачи классификации объектов недвижимости // Сети, системы связи и телекоммуникации. Деятельность вуза при переходе на федеральный образовательный стандарт 3-го поколения: материалы 34-й Всероссийской научно-технической конференции. – Рязань: Рязанское высшее военное командное училище связи, 2009. – С. 54-55.

11. Титов С.Б. Исследование проблемы управления неопределенностью при использовании fcm и rcm алгоритмов кластеризации // Математическое и программное обеспечение вычислительных систем: межвуз. сб. науч. тр. / под ред. А.Н. Пылькина. – М.: Горячая линия – Телеком, 2009. – С. 98-101.

12. Титов С.Б. Применение искусственных иммунных систем для выявления ошибочных данных в исходной выборке классифицируемых объектов недвижимости // Материалы всероссийской научной конференции «Современные исследовательские и образовательные технологии». Часть 1. – Таганрог: Изд-во ТТИ ЮФУ, 2010. – С.53-54.

13. Титов С.Б. Классификация объектов городской жилой недвижимости на основе комплексного использования инструментария искусственных иммунных систем, методов нечеткой кластеризации и теории мультимножеств // Современные проблемы информатизации в экономике и обеспечении безопасности: материалы XVI международной открытой научной конференции. – Воронеж: Издательство «Научная книга», 2011. – Вып. №16. – С. 38-41.

14. Титов С.Б. Разработка алгоритма управления степенью сжатия искусственной иммунной сети путем подбора коэффициента супрессии // Современные вопросы науки – XXI век: сборник науч. тр. по материалам VII международной научно-практической конференции. – Тамбов: Издательство ТООИПКРО, 2011. – Вып. №7. Часть 4. – С. 132-133.

15. Титов С.Б. Анализ и разработка алгоритма управления степенью сжатия искусственной иммунной сети путем подбора коэффициента супрессии при решении задачи классификации объектов городской жилой недвижимости // Математическое и программное обеспечение вычислительных систем: материалы межвузовского сборника научных трудов / под ред. А.Н. Пылькина. – Рязань (РГРТУ), 2011. – С. 111-114.

**Титов Сергей Борисович**

**ГИБРИДНЫЕ АЛГОРИТМЫ  
АНАЛИЗА И ОБРАБОТКИ ДАННЫХ  
В ЗАДАЧАХ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ**

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

Подписано в печать 11.07.2011 г. Формат бумаги 60×84 1/16.

Бумага офсетная. Печать офсетная. Усл. печ. л. 1,0.

Тираж 100 экз. Заказ

Рязанский государственный радиотехнический университет.

390005, г. Рязань, ул. Гагарина, 59/1.

Редакционно-издательский центр РГРТУ.