

На правах рукописи



Степанов Максим Анатольевич

**МОДЕЛИ ПРОГНОЗИРОВАНИЯ И АЛГОРИТМЫ АНАЛИЗА
ВРЕМЕННЫХ РЯДОВ ДЛЯ ПОДДЕРЖКИ
ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ
В СОЦИАЛЬНО-ЭКОНОМИЧЕСКОЙ СФЕРЕ**

Специальность 05.13.10 – «Управление в социальных и экономических системах»

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

Рязань – 2021

Диссертационная работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Рязанский государственный радиотехнический университет имени В.Ф. Уткина» на кафедре вычислительной и прикладной математики

Научный руководитель: **Демидова Лилия Анатольевна**
доктор технических наук, профессор,
профессор кафедры «Корпоративные
информационные системы» федерального
государственного бюджетного образовательного
учреждения высшего образования «МИРЭА –
Российский технологический университет», г. Москва

**Официальные
оппоненты:** **Гусева Анна Ивановна**
доктор технических наук, профессор,
профессор кафедры экономики и менеджмента в
промышленности федерального государственного
автономного образовательного учреждения высшего
образования «Национальный исследовательский
ядерный университет «МИФИ», г. Москва

Корнеев Андрей Матиславович
доктор технических наук, доцент,
директор института машиностроения
федерального государственного бюджетного
образовательного учреждения высшего образования
«Липецкий государственный технический
университет», г. Липецк

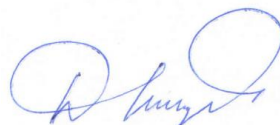
Ведущая организация: ФГБОУ ВО «Пензенский государственный
технологический университет», г. Пенза

Защита диссертации состоится «01» июля 2021 года в 12.00 на заседании диссертационного совета Д 212.211.02 в ФГБОУ ВО «РГРТУ» по адресу: 390005, г. Рязань, ул. Гагарина, д. 59/1.

С диссертацией можно ознакомиться в библиотеке ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина», а также на сайте ФГБОУ ВО «РГРТУ» www.rsreu.ru.

Автореферат разослан «__» _____ 20__ г.

Ученый секретарь
диссертационного совета Д 212.211.02,
доктор технических наук, доцент



Д.А. Перепелкин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Адекватное и обоснованное решение задач прогнозирования и анализа временных рядов (ВР) в условиях формирования системы стратегического планирования и целеполагания в настоящее время приобретает особую актуальность. Реализация Указов Президента Российской Федерации от 07.05.2018 №204 «О национальных целях и стратегических задачах развития Российской Федерации на период до 2024 года» и от 21.07.2020 № 474 «О национальных целях развития Российской Федерации на период до 2030 года», определяющих конкретное параметрическое выражение целевого уровня социально-экономического развития России на среднесрочную и долгосрочную перспективы требует от исполнителей использования комплексного инструментария анализа и прогнозирования текущего состояния и динамики значений ключевых показателей развития. Прогнозирование, в том числе на региональном уровне начинает играть существенную роль при реализации проектов и принятии управленческих решений в социально-экономической сфере.

Для оценки динамики социально-экономического развития региона в России используется более 100 показателей, представленных короткими ВР, то есть временными рядами с короткой актуальной частью, длина которой не превышает 20 – 30 отсчетов времени. Ограниченность объема актуальной информации может быть вызвана как резкими изменениями в развитии уже давно наблюдаемых ВР, так и появлением сравнительно новых ВР, для которых накопление информации только начато. При этом важной задачей остается выявление совокупности факторов, влияющих на развитие тех или иных показателей социально-экономической сферы.

Классические принципы прогнозирования ВР изложены в работах таких отечественных и зарубежных авторов, как С.А. Айвазян, Т. Андерсен, Дж. Бендит, Дж. Бокс, Д. Бриллинджер, Е.З. Демиденко, Г. Дженкинс, М. Джонсон, М. Кендэл, Б. Монтгомери, В.С. Мхитарян, А. Стьюарт, К.Чатфилд.

В число наиболее известных инструментальных средств для анализа ВР входят такие статистические пакеты, как: CART, Deductor, Forecast Expert, Predictor, SAS, SPSS, STATGRAPHICS, STATISTICA, Статистик-Консультант, ЭВРИСТА.

Однако статистические пакеты не обеспечивают эффективное решение задач анализа и прогнозирования показателей социально-экономической сферы как по причине малого объема имеющегося в наличии статистического материала, так и в виду сложности самих задач в условиях неполноты и неопределенности информации.

Нестационарность многих коротких ВР, сопоставленных социально-экономическим показателям, приводит к невозможности получения адекватных прогнозных решений с применением классических моделей и алгоритмов.

Актуальность темы исследования обусловлена существенной ролью прогнозирования в механизме управления социально-экономическим развитием регионов и при определении перспективных направлений региональной политики.

Комплексность и системность показателей, характеризующих социально-экономическое развитие регионов, делает практически невозможным обоснованное формирование ключевых положений региональной стратегии и политики без использования современных алгоритмов и моделей, обеспечивающих обоснованный и адекватный анализ и прогнозирование ВР, характеризующих показатели социально-экономической сферы.

Таким образом, существует необходимость в разработке новых подходов к решению задач анализа и прогнозирования коротких ВР, сопоставленных показателям социально-экономического развития региона. В контексте работы с показателями социально-экономической сферы, представленными короткими временными рядами, в условиях неполноты (неточности, неопределенности) и

ограниченности объемов данных и необходимости решения задач раннего выявления структурных трансформаций в группах ВР, а также – задач прогнозирования, перспективными являются технологии, реализующие применение инструментария теории нечетких множеств, генетических алгоритмов, кластерного и фазового анализа.

Основные принципы современного кластерного анализа отражены в работах таких отечественных и зарубежных авторов, как Г. Штейнгауз, С. Ллойд, Г. Болл, М. Брюинош, Д.А. Вятчинин, Р. Дженсен, Д. Дюраа, М. Жамб, Н.Г. Загоруйко, Т. Кохонен, Г. Ланс, С. МакНотон, Г. Миллиган, П. Оделл, Дж. Рубин, У. Уильямс, Х. Фридман. Алгоритмы кластерного анализа, вовлекающие в принятие решений инструментарий теории нечетких множеств, предложены и исследованы такими учеными, как Дж.К. Беждек, Дж.К. Данн, Р.Н. Дейв, Р. Кришнапурам, Я. Охаши. При этом базовым алгоритмом нечеткой кластеризации принято считать алгоритм нечетких *c*-средних (fuzzy *c*-means algorithm, FCM-алгоритм).

Ряд работ таких авторов, как Я. Батистакис, Г. Бени, Х. Галда, И. Гаф, А. Гева, Т. Калински, Л. Кауфман, П. Дж. Руссеув, С. Се, М. Халкиди, Дж. Харабаш, Ю. Фукуяма, посвящен разработке и исследованию показателей качества кластеризации.

Вопросам кластеризации ВР посвящены работы таких исследователей, как С. Агабозорги, Т.И. Вах, Д. Гада, Ж. Гама, Л. Гравано, С. Гуань, К. Калпакис, Дж. Папарризос, Ж.П. Педросо, В.Путтагунта, Ц. Пэн, П.П. Родригес, Л.Н. Феррейраа, Л. Чжаоб, Ч. Чжу, А.С. Ширхоршиди. Однако все эти работы оперируют с длинными ВР, для которых возможно применение классических статистических подходов, что принципиально не реализуемо для коротких ВР.

Аспекты фазового анализа ВР подробно отражены в работах таких отечественных и зарубежных ученых, как В.С. Анищенко, А. Вольф, К.Г. Гилмор, А.В. Головкин, В.Б. Занг, Д.П. Кратчфилд, В.А. Перепелица, Э. Петерс, Л.Н. Сергеев, Г. Шустер. При этом большинство работ связано с исследованием длинных ВР и практически не затрагивает вопросы комплексного использования полученных результатов в решении прикладных задач социально-экономической сферы.

В последние годы все большее применение при решении прикладных задач, многим из которых присуща некоторая степень неопределенности, обусловленная неполнотой (неточностью, неопределенностью) анализируемой информации, находит инструментарий теории нечетких множеств (ТНМ), базирующейся на идеях, предложенных в 1965 году Л.А. Заде.

Существенный вклад в развитие ТНМ внесли такие отечественные и зарубежные ученые, как А.Н. Аверкин, А.В. Алексеев, Р. Беллман, Л.С. Бернштейн, А.Н. Борисов, Л.А. Заде, Н.Н. Карник, А. Кофман, О.А. Крумберг, Н.Г. Малышев, Дж.М. Мендель, К. Негойце, С.А. Орловский, Р. Ягер. В настоящее время с целью минимизации вычислительных затрат обычно работают с интервальными дискретными нечеткими множествами второго типа (ИДНМТ2) на основе отпечатков неопределенности *FOU* (Footprint Of Uncertainty).

Вопросам разработки моделей прогнозирования коротких ВР с использованием инструментария ТНМ посвящены работы таких исследователей, как С.Брэд, Л.Х. Ванг, Л.В. Ли, М. Стивенсон, Дж.Э. Портер, Н. Рамли, П. Саксен, П.К.Л. Силва, К. Сонг, С.М. Чен, Й.-Ч. Ченг, Б.С. Чисс, К. Шарм, Х. Фэн, Х.-Х. Юй. При этом особое внимание уделяется разработке нечетких регрессионных моделей, что отражено в работах таких авторов, как М.А. Вида, Х. Ли, Н. Марон, Ш. Педрич, Д.А. Савич, М. Сакава, Дж.К. Сонг, Х. Танака, Д.Х. Хонг, Х. Яно.

Отечественные основы управления социально-экономическими системами (СЭС) разработаны такими учеными, как: А.Г. Гранберг, Н.Н. Колосовский, С.М. Лавлинский, А.А. Мясников, Н.Н. Некрасов, Д.А. Новиков и др.

Как показывает анализ результатов прогнозирования с применением моделей, предполагающих привлечение инструментария ТНМ, довольно часто такие модели не обеспечивают получение адекватных результатов прогнозирования по причине недостаточно обоснованного выбора значений параметров моделей. Решение выявленной проблемы может быть получено с привлечением тех или иных эволюционных алгоритмов оптимизации.

Эволюционные алгоритмы оптимизации обеспечивают адекватное решение прикладных задач, трудноразрешимых классическими методами оптимизации (в том числе, – *NP*-полных задач оптимизации). Наиболее известными эволюционными алгоритмами оптимизации являются генетические алгоритмы (ГА), реализующие поиск оптимальных решений с привлечением принципов эволюционных вычислений, основанных на генетических процессах, происходящих в биологических организмах. При этом осуществляется поиск экстремума некоторой целевой функции – так называемой функции соответствия (*fitness function*). Базовые принципы ГА были сформулированы в 1975 году Дж.Х. Холландом и получили свое развитие в работах таких ученых, как Н.А. Барричелли, Д.И. Батищев, Л.А. Гладков, Д.И. Голдберг, В.В. Емельянов, А.П. Карпенко, В.В. Курейчик, В.М. Курейчик, А.П. Ротштейн, Д. Рутковская, Е.С. Семенкин, Л.Дж. Фогель. Одним из важных достижений в области эволюционных алгоритмов оптимизации следует считать адаптацию этих алгоритмов к решению многокритериальных задач оптимизации. При этом одним из эффективных считается многокритериальный генетический алгоритм оптимизации, предложенный группой ученых во главе с К. Дебом в 2002 году и реализующий формирование фронта Парето-оптимальных решений.

Использование принципов многокритериальной оптимизации при реализации генетических алгоритмов должно обеспечить решение проблемы выбора оптимальных значений параметров моделей прогнозирования коротких ВР, предполагающих привлечение инструментария теории нечетких множеств, при приемлемых временных затратах.

Таким образом, несмотря на наличие значительного числа работ, посвященных проблемам прогнозирования ВР, вопросам кластерного и фазового анализа, аспектам теории нечетких множеств и эволюционных вычислений, существует необходимость разработки новых подходов к анализу и прогнозированию ВР социально-экономической сферы, с целью повышения качества информации для разработки адекватной и обоснованной региональной политики.

Цель работы – повышение обоснованности и адекватности принимаемых управленческих решений посредством разработки моделей прогнозирования и алгоритмов анализа показателей социально-экономической сферы региона, актуальные значения которых представлены короткими временными рядами, и автоматизации указанных процессов прогнозирования и анализа.

Для достижения поставленной цели решаются следующие задачи.

1 Выполнить анализ применимости инструментария теории нечетких множеств и генетических алгоритмов для разработки регрессионных и многокритериальных моделей прогнозирования коротких временных рядов, исследовать возможности инструментария кластерного и фазового анализа в задаче выявления структурных трансформаций в группах временных рядов.

2 Разработать модели прогнозирования и анализа временных рядов с использованием инструментария теории нечетких множеств и генетических алгоритмов.

3 Разработать методику оценки вариативности коротких временных рядов социально-экономической сферы.

4 Разработать алгоритмы выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы с применением инструментария кластерного и фазового анализа.

5 Разработать пакет прикладных программ (ППП) для автоматизации анализа и прогнозирования коротких ВР с применением предложенных алгоритмов и моделей.

Объектом исследования являются технологии анализа и прогнозирования временных рядов.

Предметом исследования являются показатели социально-экономической сферы региона, актуальные значения которых представлены короткими временными рядами.

Область исследования соответствует паспорту научной специальности 05.13.10 – Управление в социальных и экономических системах, пункту 6 «Разработка и совершенствование методов получения и обработки информации для задач управления социальными и экономическими системами», пункту 10 «Разработка методов и алгоритмов интеллектуальной поддержки принятия управленческих решений в экономических и социальных системах».

Научная новизна результатов диссертационного исследования.

1 Предложена модель прогнозирования и анализа коротких временных рядов на основе регрессионного анализа, отличающаяся использованием уравнения нечеткой регрессии с асимметричными параметрами и генетического алгоритма, позволяющая оценить уровень неопределенности поведения временного ряда в контексте выбранного уравнения нечеткой регрессии, а также сформировать на его основе сценарный прогноз показателей социально-экономической сферы региона на один шаг вперед.

2 Предложена универсальная многокритериальная *FOU*-модель прогнозирования, отличающаяся использованием двух минимизируемых критериев, позволяющая посредством реализации генетического алгоритма сделать сценарный прогноз на один шаг вперед и охватывающая частный случай, предполагающий представление временных рядов на основе дискретных нечетких множеств первого типа.

3 Предложена методика оценки вариативности коротких временных рядов социально-экономической сферы, отличающаяся выполнением двухэтапного анализа временного ряда с учетом значений элементов как исходного временного ряда, так и временного ряда, содержащего значения показателя Херста или фрактальной размерности, с применением универсальной многокритериальной *FOU*-модели прогнозирования и позволяющая повысить обоснованность оценки вариативности временного ряда.

4 Разработан алгоритм выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы, отличающийся использованием нечетких *s*-средних и позволяющий, в случае хорошей отделимости кластеров, выполнить анализ поведения групп показателей социально-экономического развития и выявить структурные трансформации (изменения в их динамике) на ранних стадиях появления.

5 Разработаны алгоритмы выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы, отличающиеся использованием иерархической агломеративной кластеризации, не накладывающей на группу временных рядов жесткие требования хорошей отделимости кластеров, и позволяющие выполнить анализ поведения групп показателей социально-экономического развития и выявить структурные трансформации на ранних стадиях

появления в случаях, когда осуществляется анализ показателей социально-экономического развития, относящихся к разным подсистемам социально-экономической системы.

Теоретическая и практическая значимость работы. Основные положения диссертационной работы вносят вклад в развитие методов обработки информации для поддержки принятия управленческих решений, в частности, в развитие средств прогнозирования и анализа показателей социально-экономической сферы, представленных короткими временными рядами, – предложены и исследованы новые модели прогнозирования и алгоритмы анализа указанных коротких временных рядов на основе инструментария теории нечетких множеств, генетических алгоритмов и кластерного анализа.

Выводы и результаты исследования ориентированы на практическое применение предложенных моделей и алгоритмов, использованы при анализе инвестиционных процессов и общего состояния экономического развития Рязанской области. Кроме того, результаты могут быть использованы региональными органами исполнительной власти при формировании краткосрочного сценарного прогноза социально-экономического развития и анализе тенденций регионального социально-экономического развития с целью повышения качества управленческих решений.

Самостоятельную практическую значимость имеет разработанный ППП, реализующий анализ и прогнозирование коротких ВР социально-экономической сферы с применением предложенных алгоритмов и моделей, подтверждающий их адекватность и эффективность.

Методы исследования. В диссертационной работе использовались методы теории нечетких множеств, теории вероятностей, системного анализа, кластерного анализа, фазового анализа, математической статистики, интеллектуальной обработки информации, эволюционного моделирования, оптимизации, модульного и объектно-ориентированного программирования.

Основные положения, выносимые на защиту:

1 Модель прогнозирования и анализа коротких временных рядов социально-экономической сферы с использованием регрессионной модели прогнозирования на основе уравнения нечеткой регрессии с асимметричными параметрами и генетического алгоритма, позволяющая оценить уровень неопределенности поведения временного в контексте выбранного уравнения нечеткой регрессии, а также сформировать сценарный прогноз на один шаг вперед, учитывающий возможные экстремальные отклонения значений показателей.

2 Универсальная многокритериальная *FOU*-модель прогнозирования, позволяющая посредством реализации генетического алгоритма обеспечить высокое качество результатов прогнозирования на один шаг вперед с формированием сценарного прогноза при приемлемых временных затратах, снижающая в ряде случаев ошибку прогнозирования показателей социально-экономической сферы региона на 0,9 – 2,1 процентных пункта.

3 Методика оценки вариативности временного ряда социально-экономической сферы, предполагающая выполнение двухэтапного анализа временного ряда с учетом значений элементов как исходного временного ряда, так и временного ряда, содержащего значения показателя Херста или фрактальной размерности, с применением универсальной многокритериальной *FOU*-модели прогнозирования, позволяющая повысить обоснованность оценки вариативности временного ряда.

4 Алгоритмы выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы на основе алгоритмов нечеткой и иерархической агломеративной кластеризации, реализующие адекватное и обоснованное разбиение групп временных рядов на кластеры для различных длин

временного ряда, получаемых при увеличении числа наблюдаемых отсчетов времени, и обеспечивающие раннее выявление структурных трансформаций в группах временных рядов при приемлемых временных затратах посредством сравнительного анализа результатов кластеризации для каждого двух последовательных отсчетов времени.

Степень достоверности и апробация результатов. Достоверность полученных результатов подтверждается корректным использованием методов теории нечетких множеств, теории вероятностей, системного анализа, математической статистики, интеллектуальной обработки информации, эволюционного моделирования, оптимизации, модульного и объектно-ориентированного программирования, тестированием разработанных алгоритмов и моделей, сравнением полученных результатов с результатами, полученными в специализированных пакетах прикладных программ, реализующих анализ временных рядов.

Исследования по тематике диссертационной работы проводились в рамках следующих научных проектов, поддержанных Российским фондом фундаментальных исследований (РФФИ):

- проект № 18-410-623001 «Разработка механизмов и методов диагностики и прогнозирования структурных трансформаций экономики региона с учетом приоритетов устойчивого социально-экономического развития» (2018-2020 гг.) – руководитель проекта;

- проект № 20-010-00219 «Оценка влияния цифровизации на развитие промышленного производства» (2020 – 2021 гг.) – исполнитель проекта.

Основные положения и результаты диссертационной работы докладывались и обсуждались на 6 Всероссийских и 7 Международных конференциях: XXII, XXIII, XXIV Всероссийских научно-технических конференциях студентов, молодых ученых и специалистов «Новые информационные технологии в научных исследованиях и в образовании (НИТ-2018, НИТ-2019; НИТ-2020)» (г. Рязань, 2018, 2019, 2020); III Всероссийской научно-технической конференции «Актуальные проблемы современной науки и производства» (г. Рязань, 2018); Всероссийской научно-технической конференции «Интеллектуальные и информационные системы (Интеллект-2019)» (г. Тула, 2019); Международной научно-технической и научно-методической конференции «Современные технологии в науке и образовании (СТНО-2017, СТНО-2019, СТНО-2020)» (г. Рязань, 2017, 2019, 2020); Международной научно-технической конференции «Актуальные проблемы прикладной математики, информатики и механики» (г. Воронеж, 2019); 1st, 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA2019, SUMMA2020) (г. Липецк, 2019, 2020); 9th, 10th Mediterranean Conference on Embedded Computing (Budva, Montenegro, 2019, 2020); 2020 IEEE International Conference on Information Technologies (InfoTech-2020 34th issue) (Varna, Bulgaria).

Алгоритмы и модели, разработанные в диссертационной работе, а также реализующий их ППП внедрены и используются в следующих организациях:

1) АО «Корпорация развития Рязанской области» при анализе тенденций социально-экономического развития Рязанской области, в частности анализа инвестиционных процессов (акт внедрения от 18.08.2020).

2) ООО «Сафьян» при прогнозировании показателей финансово-хозяйственной деятельности организации в качестве дополнительного источника аналитической информации (акт внедрения от 14.04.2020).

3) ФГБОУ ВО «Рязанский государственный радиотехнический университет им. В.Ф. Уткина» в учебный процесс по направлениям подготовки «Прикладная информатика» и «Программная инженерия» при проведении лекций, лабораторных

работ и практических занятий по дисциплинам «Теория принятия решений», «Методы представления и обработки данных», «Методы интеллектуального анализа данных», «Эволюционные методы оптимизации», а также при курсовом, дипломном проектировании и в студенческой НИР (акт внедрения от 30.06.2020).

Получено 2 свидетельства ФГБУ «Федеральный институт промышленной собственности» Федеральной службы по интеллектуальной собственности (РОСПАТЕНТ) о государственной регистрации программ для ЭВМ.

Публикации. Всего по теме диссертации опубликовано 23 работы, в том числе: 4 статьи в ведущих рецензируемых научных журналах из Перечня ВАК, 8 работ проиндексированы в международной базе данных Scopus, 11 тезисов докладов на Международных и Всероссийских научно-технических конференциях; получено 2 свидетельства о государственной регистрации программ для ЭВМ (Приложение 6).

Структура и объем диссертации. Диссертация состоит из введения, четырех глав, списка использованных источников. Содержит 208 страниц без библиографического списка и приложений, 13 таблиц, 54 рисунка. Библиографический список состоит из 178 наименований. В приложениях приведены результаты экспериментальных исследований в графическом и табличном виде (Приложения 1 – 4), акты о внедрении результатов диссертационной работы (Приложение 5) и свидетельства о государственной регистрации программ для ЭВМ (Приложение 6).

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, определены объект и предмет исследования. Сформулированы цель и задачи исследования, основные положения, выносимые на защиту. Показана научная новизна и практическая значимость основных научных результатов. Дана общая характеристика работы.

В первой главе выполнен анализ применимости инструментария теории нечетких множеств и генетических алгоритмов для разработки регрессионных и многокритериальных моделей прогнозирования коротких временных рядов, исследованы возможности инструментария кластерного и фазового анализа в задаче выявления структурных трансформаций в группах временных рядов.

1 Выполнена постановка задач прогнозирования и анализа ВР социально-экономической сферы для расширения объема исходной информации при принятии управленческих решений. Выявлены характерные проблемы, вызванные ограниченностью объемов актуальных данных. Показана актуальность работы.

2 Исследован подход к прогнозированию и анализу ВР, реализующий применение инструментария теории нечетких множеств (ТНМ). Высказано предположение о перспективности использования инструментария ТНМ при решении задач разработки моделей прогнозирования ВР, в том числе – коротких ВР.

3 Высказаны предположения о целесообразности выполнения анализа фазовых портретов ВР, а также о перспективности прогнозирования и анализа ВР, на основе построения уравнений нечетких регрессий.

4 Сделано предположение о целесообразности использования генетических алгоритмов (ГА) для поиска оптимальных значений параметров моделей прогнозирования на основе нечетких множеств, а также оптимальных значений параметров уравнений нечетких регрессий.

5 Рассмотрены общие аспекты, изложены принципы и показана специфика задачи кластеризации ВР. Высказано предположение о перспективности использования нечетких и иерархических алгоритмов кластеризации в контексте решения задачи кластеризации коротких ВР с целью выявления структурных трансформаций в группах ВР социально-экономической сферы.

На основании анализа уточнены задачи исследования. Повышение обоснованности и адекватности принимаемых управленческих решений предлагается осуществить посредством создания научно-обоснованного инструментария, позволяющего оперировать информацией о реальном состоянии социально-экономической сферы региона, полученной с привлечением современного математического аппарата и подтвержденной с применением соответствующих критериев (показателей) качества решений.

Во второй главе предложены и исследованы алгоритмы выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы с использованием инструментария нечеткой и иерархической кластеризации, предложена модель прогнозирования и анализа коротких временных рядов на основе регрессионного анализа, отличающаяся использованием уравнения нечеткой регрессии с асимметричными параметрами.

1 Разработан алгоритм выявления структурных трансформаций в группах коротких ВР социально-экономической сферы с использованием нечетких c -средних.

Этот алгоритм включает следующую последовательность шагов.

Шаг 1. Реализовать кластеризацию группы ВР длиной $(t - 1)$ и группы ВР длиной t с применением FCM -алгоритма с целью формирования кластеров ВР, обладающих сходной динамикой развития, используя алгоритм нечетких c -средних для выявления центроидов кластеров (усредненных ВР) при фиксированном значении числа c кластеров и использования их для прогнозирования динамики поведения группы ВР, кластеров группы ВР, а также отдельных ВР в кластерах.

Шаг 2. Выполнить сравнительный анализ оптимальных чисел $c_{(t-1)}$ и c_t кластеров для групп ВР длиной $(t - 1)$ и t . Если $c_{(t-1)} = c_t$ принять решение об отсутствии больших структурных трансформаций в группе ВР. В противном случае принять решение о наличии больших структурных трансформаций в группе ВР.

Если $c_{(t-1)} < c_t$, выявить, какие ВР сформировали новый кластер (кластеры). Определить, какие кластеры были удалены, если $c_{(t-1)} > c_t$. Выполнить сравнительный анализ содержимого кластеров для группы ВР длиной $(t - 1)$ и группы ВР длиной t .

Если $c_{(t-1)} = c_t$, выполнить сравнительный анализ содержимого кластеров для группы ВР длиной $(t - 1)$ и группы ВР длиной t . Если содержимое кластеров изменилось, принять решение о наличии средних структурных трансформаций в группе ВР. В противном случае принять решение об отсутствии средних структурных трансформаций в группе ВР.

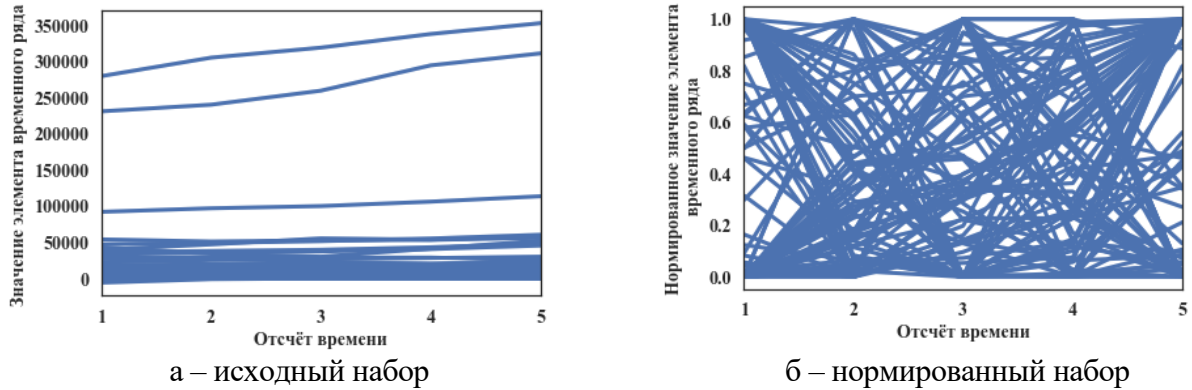
Если $c_{(t-1)} = c_t$ и средние структурные трансформации отсутствуют, осуществить проверку содержимого кластеров на наличие малых структурных трансформаций, выполнив упорядочение ВР внутри кластеров по убыванию расстояний от центроида каждого кластера до каждого ВР этого кластера. Если списки упорядочения не изменились, принять решение об отсутствии малых структурных трансформаций в группе ВР. В противном случае принять решение о наличии малых структурных трансформаций в группе ВР. Перейти к шагу 3.

Шаг 3. Осуществить принятие решений по результатам анализа. В случае плохой отделимости кластеров принять решение о целесообразности исследования с применением иерархической кластеризации. Завершить работу алгоритма.

Предлагаемый алгоритм позволяет выявлять на ранней стадии возможные проблемы (в частности, структурные трансформации) в динамике ВР и предпринимать адекватные меры по их устранению, а также прогнозировать наличие/отсутствие положительного эффекта от предпринимаемых мер.

Апробация предлагаемого алгоритма была выполнена на примере реальной группы (набора) показателей социально-экономического развития Рязанской области за 7 лет (с 2011 по 2017 годы), в которой насчитывалось 89 показателей,

представленных короткими ВР (длиной не более 7 отсчетов времени) и сгруппированных по 9 категориям: «Население», «Производство товаров и услуг», «Торговля и услуги населению», «Внеэкономическая деятельность», «Малое и среднее предпринимательство, включая микропредприятия», «Инвестиции», «Консолидированный бюджет субъекта Российской Федерации (включая местные бюджеты без учета территориальных внебюджетных фондов)», «Денежные доходы и расходы населения» и «Труд и занятость» (рисунок 1).



а – исходный набор

б – нормированный набор

Рисунок 1 – Визуализация реального набора ВР

Несмотря на недостаточно хорошую делимость кластеров друг от друга в реальном наборе ВР, применение *FCM*-алгоритма при числе отсчетов, равном 3, 4 и 5, позволило выделить возможное оптимальное число кластеров: оно оказалось равным соответственно 3, 12 и 10. Эти результаты были получены с использованием в качестве показателя качества кластеризации индекса C_e – Бени (Xie-Beni index). Получение существенно разных значений для оптимального числа кластеров при разном числе отсчетов времени позволяет сделать вывод о наличии больших структурных трансформаций в группе ВР.

Для категории ВР «Производство товаров и услуг», насчитывающей в своем составе 15 ВР, длина которых равна 7 отсчетам времени, применение *FCM*-алгоритма при числе отсчетов, равном 3, 4, 5, 6 и 7, позволило выделить возможное оптимальное число кластеров: оно оказалось равным соответственно 4, 9, 7, 10 и 10. Получение разных значений для оптимального числа кластеров при числе отсчетов времени, равном 3, 4, 5 и 6, позволяет сделать вывод, что при увеличении числа отсчетов на 1 каждый раз имеет место наличие больших структурных трансформаций в группе ВР.

Кластеры, содержащиеся в рассматриваемой группе ВР, плохо делимы друг от друга, а выявленное с применением *FCM*-алгоритма оптимальное число кластеров не может быть признано удовлетворительным для дальнейшего решения задачи выявления структурных трансформаций ВР.

2 Разработаны алгоритмы выявления больших, средних и малых структурных трансформаций в группах коротких временных рядов социально-экономической сферы, с использованием иерархической агломеративной кластеризации, не накладывающей на группу ВР жесткие требования хорошей делимости кластеров.

Предложенные алгоритмы по принципу реализации аналогичны алгоритму выявления структурных трансформации на основе нечетких *c*-средних. Эти алгоритмы реализуются последовательно. При этом в качестве метода слияния кластеров используется метод Уорда, а в качестве метрики для вычисления расстояния между объектами – евклидова метрика, позволившие обеспечить наилучшие результаты кластеризации с точки зрения применяемых показателей качества кластеризации (индекс кластерного силуэта и *elbow/knee*-индекс).

Принятие решений о наличии структурных трансформаций выполняется в соответствии со следующими особенностями.

Большие структурные трансформации в группах ВР присутствуют в случае, если при переходе от ВР длиной $(t - 1)$ к ВР длиной t , число кластеров, полученных при разбиении группы ВР, не совпадает.

Средние структурные трансформации в группах ВР присутствуют в случае, если при переходе от ВР длиной $(t - 1)$ к ВР длиной t , число кластеров, полученное при разбиении группы ВР длиной $(t - 1)$, совпадает с числом кластеров, полученным при разбиении группы ВР длиной t , но при этом наблюдается изменение структуры существующих кластеров.

Средние структурные трансформации в группах ВР присутствуют в случае, если при переходе от ВР длиной $(t - 1)$ к ВР длиной t , число кластеров, полученное при разбиении группы ВР длиной $(t - 1)$, совпадает с числом кластеров, полученным при разбиении группы ВР длиной t , структура существующих кластеров не изменяется, но имеет место изменение порядка включения ВР в свои кластеры.

На рисунке 2 приведен пример иерархической кластеризации нормированной группы ВР для 89 показателей социально-экономического развития Рязанской области при длине ВР в 5 элемента. В результате, группа ВР может быть разбита на 3 кластера, при этом, при изменении длины ВР различны составы кластеров, в связи с чем можно сделать вывод о наличии средних структурных трансформаций в группе ВР в целом.

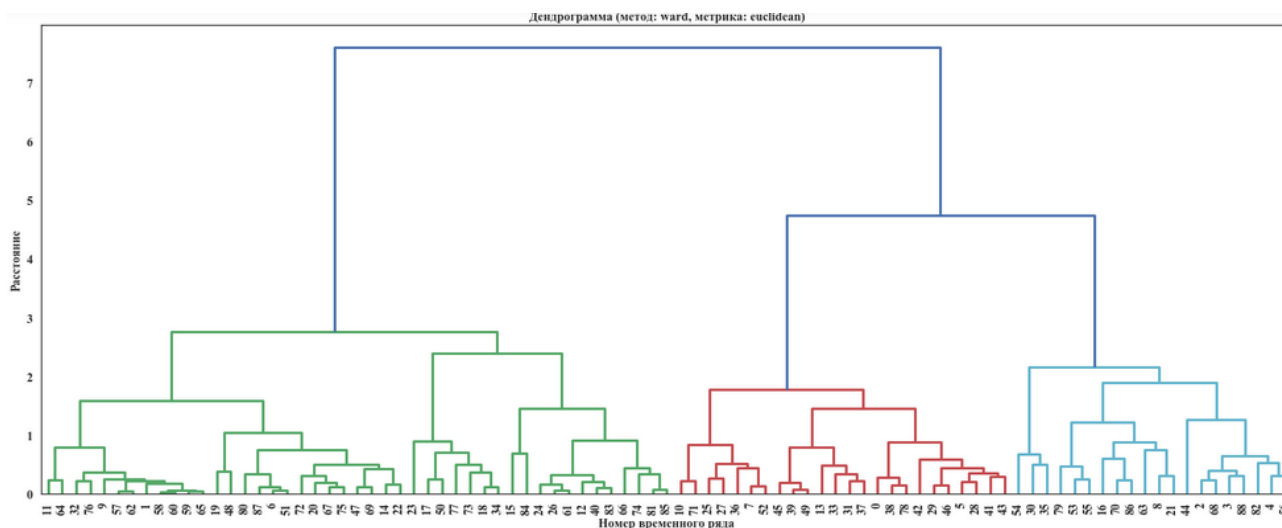


Рисунок 2 – Дендрограмма для нормированной группы ВР длиной в 5 элементов

Предлагаемые алгоритмы были применены к группам ВР, относящимся к отдельным категориям. Визуализация результатов по категории «Инвестиции», содержащей 13 ВР при числе отсчетов времени, равном 4 и 5, представлена на рисунке 3, выявлены средние структурные трансформации.

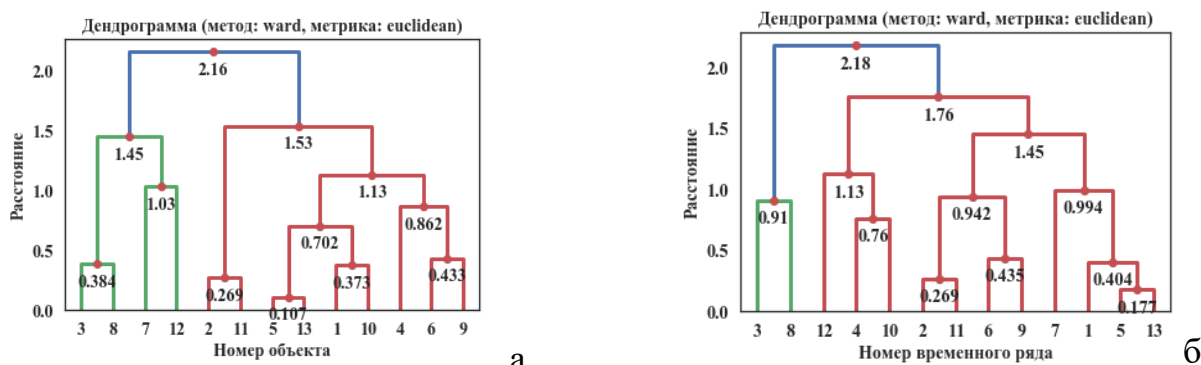


Рисунок 3 – Дендрограммы ВР групп категории «Инвестиции»:

а – дендрограмма нормированной группы ВР категории «Инвестиции» при $t = 4$

б – дендрограмма нормированной группы ВР категории «Инвестиции» при $t = 5$

Это связано с тем, что показатели категории «Инвестиции» во многом определяются инвестиционным климатом в регионе, формирование которого происходит как в долгосрочной, так и в среднесрочной перспективе при реализации инвестиционной политики.

Результаты анализа социально-экономической сферы Рязанской области с использованием предлагаемых алгоритмов позволили выявить ряд структурных трансформаций в группах ВР (таблица 1).

Таблица 1 – Структурные трансформации групп ВР социально-экономической сферы Рязанской области

Категория структурных трансформаций	Категория ВР	Характеристика и возможные меры управленческого реагирования
Большие	- Труд и занятость	Наличие больших структурных трансформаций по категории ВР «Труд и занятость» в рязанской области связано с поэтапным выходом данной сферы из кризисной ситуации. При общем снижении безработицы наблюдается устойчивый рост среднемесячной заработной платы. Таким образом, подтверждается эффективность проводимой политики по поддержке занятости.
Средние	- Валовой региональный продукт - Промышленное производство - Строительство - ВЭД - Инвестиции	Средние структурные трансформации характерны для всех групп ВР, характеризующих состояние экономики, производства и экономической активности и свидетельствует о нарастании кризисных явлений в экономике региона и характеризуемых, прежде всего, разнонаправленными колебаниями показателей и индексов. Наличие структурных трансформаций свидетельствует о том, что колебания вызваны не просто общеэкономической кризисной ситуацией, и выдвигает требования о необходимости формирования новых подходов к реализации региональной экономической политики.
Малые	- Консолидированный бюджет субъекта Российской Федерации	Малые структурные трансформации по категории ВР, связанных с региональным бюджетом, вызваны ростом расходной части бюджета при незначительно меняющейся доходной.

Предлагаемые алгоритмы выявления структурных трансформаций в группах ВР основаны на современной инструментари кластерного анализа и позволяют на основе раннего выявления структурных трансформаций в группах ВР социально-экономической сферы региона идентифицировать и предотвращать развитие кризисных явлений, а также формировать приоритеты региональной политики.

Так, полученная информация о наличии структурных трансформаций в группах ВР социально-экономической сферы Рязанской области (в частности по группам «Инвестиции», «Промышленное производство», «Труд и занятость») стала основой при поддержке принятия решений о выборе приоритетной сферы и инструментов инвестиционного развития (акт внедрения АО «Корпорация развития Рязанской области» от 18.08.2020).

3 Предложена методика выявления актуальной длины короткого ВР на основе анализа фазовых портретов.

Анализ фазового портрета ВР, содержащего m элементов, позволяет выявить все присутствующие в анализируемом ВР аттракторы, оценить их длину и ограничить длину ВР, используемого при разработке модели прогнозирования, в целях выявления наиболее актуальной информации.

Методика выявления актуальной длины анализируемого ВР может быть представлена следующей последовательностью шагов:

- построить фазовый портрет анализируемого ВР;

- выявить все присутствующие в анализируемом ВР аттракторы;
- определить актуальную длину анализируемого ВР.

Построение фазового портрета анализируемого ВР предлагается реализовать с использованием известного подхода, предполагающего, что по оси абсцисс должно откладываться значение d_i ($i = \overline{2, m}$) элемента ВР в каждый текущий отсчет времени, а по оси ординат – его цепной абсолютный прирост $\Delta d_i = d_i - d_{i-1}$ ($i = \overline{2, m}$).

Анализ фазовых портретов одновременно развивающихся взаимосвязанных ВР, позволяет выявить присутствующие в них аттракторы, оценить их длину и предсказать появление структурных трансформаций в одних ВР, выявив структурные трансформации в других ВР (если известно, что некоторые ВР схоже реагируют на одинаковые внешние воздействия с определенным лагом).

Предлагаемая методика выявления актуальной длины анализируемого ВР позволяет учитывать при разработке моделей прогнозирования действительно актуальные элементы ВР.

4 Предложена модель прогнозирования и анализа коротких временных рядов с использованием уравнения нечеткой регрессии с асимметричными параметрами и генетического алгоритма, позволяющая сформировать сценарный прогноз показателей социально-экономической сферы региона на один шаг вперед.

Использование ГА для подбора оптимальных значений параметров модели прогнозирования, а также введение асимметричности для значений параметров позволяют сформировать сценарный прогноз, адекватный реальному состоянию социально-экономической сферы региона, и, как результат, повысить обоснованность управленческих решений, принимаемых с учетом такого прогноза.

Описание поведения короткого ВР может быть выполнено на основе модели нечеткой регрессии (НР). В качестве уравнений регрессии предлагается использовать уравнения линейной (НЛР), квадратичной (НКР), экспоненциальной (НЭР) и логарифмической (НЛогР) нечетких регрессий, в которых вместо четких значений параметров используются их нечеткие аналоги, представленные с помощью асимметричных треугольных нечетких чисел (ТНЧ, Triangle Fuzzy Number, TFN) вида $TFN = (e, e_-, e_+)$, где e – точка универсума, в которой достигается максимальное значение функции принадлежности TFN , равное 1, а e_- и e_+ – точки универсума, в которых значения функции принадлежности TFN равны 0 ($e_- < e$; $e_+ > e$).

Уравнения нечетких регрессий, соответствующие одноименным моделям прогнозирования – могут быть записаны в виде:

$$Y(x) = TFN_a \cdot x + TFN_b; \quad (1)$$

$$Y(x) = TFN_a \cdot x^2 + TFN_b \cdot x + TFN_c; \quad (2)$$

$$Y(x) = TFN_a \cdot \exp(TFN_b \cdot x) + TFN_c; \quad (3)$$

$$Y(x) = TFN_a \cdot \log(TFN_b + x) + TFN_c. \quad (4)$$

Для определения значений параметров уравнения нечеткой регрессии составляется и решается задача квадратичного программирования (ЗКП), в результате чего определяются искомые $TFN_a = (a_-, a, a_+)$ и $TFN_b = (b_-, b, b_+)$.

Из-за асимметричности ТНЧ $TFN_a = (a_-, a, a_+)$ и $TFN_b = (b_-, b, b_+)$ коридор уравнения НЛР анализируемого ВР также асимметричен.

Для точек анализируемого ВР и ВР, соответствующего уравнению классической регрессии, значение нечеткой меры подобия может быть вычислено, например, как:

$$f = 1 - \frac{\sum_{j=1}^m (1 - u_{TS}(x_j, y'_j))}{\sum_{j=1}^m (1 + u_{TS}(x_j, y'_j))}, \quad (5)$$

где $u_{TS}(x_j, y'_j)$ – значение функции принадлежности нечеткого множества TS (Time Series) анализируемого ВР к уравнению КР для НР этого ВР для значения элемента ВР y'_j в отсчет времени x_j ($j = \overline{1, m}$); m – число элементов ВР.

Используя уравнение НР для анализируемого ВР, можно сформировать пессимистическое и оптимистическое прогнозные решения, опираясь на нижнюю и верхнюю границы коридора НР, а также нейтральное (среднее, классическое) прогнозное решение. При работе с уравнением нечеткой регрессии с асимметричными параметрами, пессимистическое и оптимистическое прогнозные решения будут иметь разные значения отклонений от нейтрального (среднего, классического) прогнозного решения.

Формула для вычисления значения функция принадлежности (ФП) точки (x', y') к уравнению НР может быть записана как:

$$u(x', y') = \begin{cases} 1 - \frac{f(x'; a, b, c) - y'}{f(x'; a_-, b_-, c_-)}, & \text{если} \\ f(x'; a, b, c) - f(x'; a_-, b_-, c_-) \leq y' \leq f(x'; a, b, c); \\ 1 - \frac{y' - f(x'; a, b, c)}{f(x'; a_+, b_+, c_+)}, & \text{если} \\ f(x'; a, b, c) \leq y' \leq f(x'; a, b, c) + f(x'; a_+, b_+, c_+); \\ 0, & \text{иначе,} \end{cases} \quad (6)$$

где $f(x'; a, b, c)$, $f(x'; a_-, b_-, c_-)$ и $f(x'; a_+, b_+, c_+)$ – теоретические значения, полученные с помощью уравнения классической регрессии для нечеткой регрессии того или иного типа при отсчете времени x' и соответствующих значениях параметров ТНЧ ($a, b, c, a_-, b_-, c_-, a_+, b_+, c_+$); y' – реальное значение элемента ВР при отсчете времени x' .

Для обеспечения одновременного поиска типа уравнения нечеткой регрессии и значений его параметров при приемлемых временных затратах предлагается использовать ГА, в котором в качестве функции соответствия используется нечеткая мера подобия, вычисляемая, например, по формуле (5), и подлежащая максимизации.

Предложенный ГА может быть описан последовательностью шагов:

Шаг 1. Создать начальную популяцию хромосом вида размером P , определив случайным образом значения генов в границах допустимых диапазонов: $type \in \{1, 2, 3, 4\}$; $|a| \leq A$; $|b| \leq B$; $|c| \leq C$; $a_- \geq 0$; $b_- \geq 0$; $c_- \geq 0$; $a_+ \geq 0$; $b_+ \geq 0$; $c_+ \geq 0$; $a_- \leq A_-$; $b_- \leq B_-$; $c_- \leq C_-$; $a_+ \leq A_+$; $b_+ \leq B_+$; $c_+ \leq C_+$, где $A; B; C$; $A_-; B_-; C_-$; $A_+; B_+; C_+$ – числа, задающие границы соответствующих диапазонов. Положить текущее число поколений g равным 1.

Шаг 2. Вычислить значения функции соответствия на основе нечеткой меры подобия для всех хромосом популяции и упорядочить их по убыванию вычисленных значений.

Шаг 3. Сформировать пары хромосом-родителей, используя метод рулетки, при реализации которого вероятность выбора i -й хромосомы ($i = \overline{1, P}$) в качестве родителя

вычисляется как $p_i = f_i / \sum_{i=1}^P f_i$, где f_i – значение функции соответствия. Вероятность p_i тем больше, чем больше у i -й хромосомы значение функции соответствия f_i .

Шаг 4. Выполнить для каждой пары хромосом генерацию случайного равномерно распределенного числа z_c из отрезка $[0, 1]$. Если $z_c < Z_c$, выполнить операцию скрещивания хромосом-родителей, сгенерировав случайное равномерно распределенное число $N_c \in \{1, 2, 3\}$, определяющее точку скрещивания.

Шаг 5. Выполнить для хромосом, выбранных в качестве хромосом-родителей на шаге 4, генерацию случайного равномерно распределенного числа z_m из отрезка $[0, 1]$. Если $z_m < Z_m$, выполнить операцию мутации хромосом-родителей, сгенерировав случайное равномерно распределенное числа $N_m \in \{1, 2, 3, 4\}$, определяющее точку мутации. Вычислить значения функции соответствия для хромосом-потомков.

Шаг 6. Сформировать расширенную популяцию размером $P + Z_c \cdot P$ из хромосом текущего поколения и хромосом-потомков, полученных на шагах 4 и 5. Упорядочить расширенную популяцию хромосом по убыванию значений функции соответствия. Удалить $Z_c \cdot P$ хромосом с худшими (наименьшими) значениями функции соответствия из расширенной популяции для получения популяции размером P .

Шаг 7. Перейти к шагу 3, если текущее число поколений g меньше, чем максимальное число поколений G . Иначе увеличить текущее число поколений g на 1 и перейти к шагу 8.

Шаг 8. Принять в качестве искомого решение, закодированное хромосомой с максимальным значением функции соответствия.

При этом реализуется принцип элитизма: лучшие решения передаются из текущего поколения в последующее.

При реализации ГА хромосома может быть закодирована в виде:

$$ch = (type, a, a_-, a_+, b, b_-, b_+, c, c_-, c_+), \quad (7)$$

где $type$ – ген, отвечающий за тип уравнения нечеткой регрессии и принимающий значения 1, 2, 3 и 4 для уравнений линейной, квадратичной, экспоненциальной и логарифмической регрессий.

Реализация ГА позволила одновременно найти тип уравнений нечеткой регрессии (один из четырех), а также значения параметров треугольных нечетких чисел при приемлемых временных затратах.

Построение коридоров уравнения нечеткой регрессии было использовано при формировании прогнозов на один шаг вперед. На основе знаний о нижней и верхней границах коридора нечеткой регрессии были сформированы оптимистическое и пессимистическое прогнозные решения, а также – нейтральное прогнозное решение.

На рисунке 4 представлен пример прогнозирования на один шаг вперед для ВР, характеризующего показатель развития Рязанской области «Объем промышленного производства».

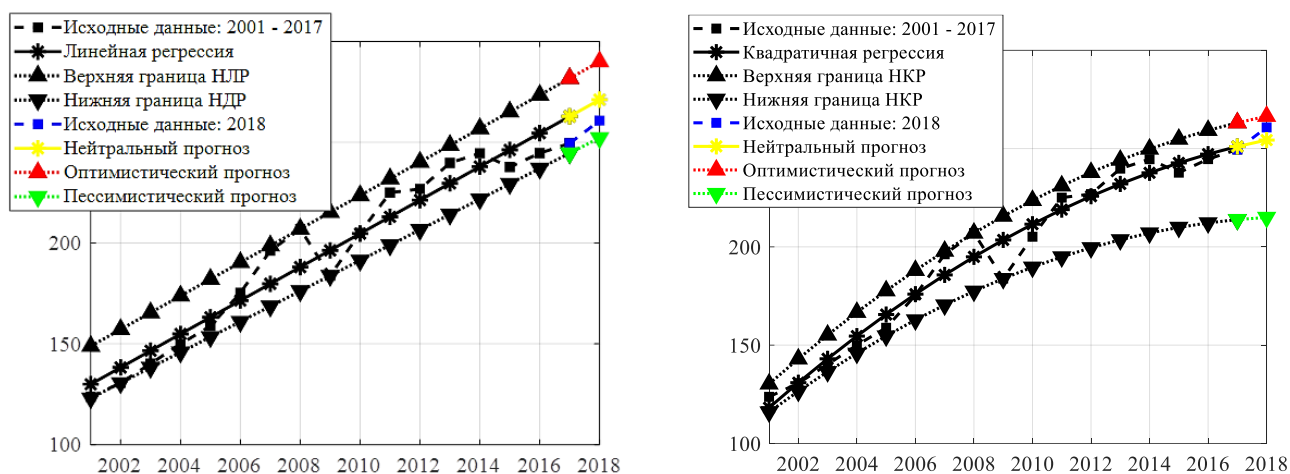


Рисунок 4 – Коридор НКР и прогноз на один шаг вперед для показателя «Объем промышленного производства» с использованием уравнений НЛР и НКР

В качестве анализируемой длины ВР была выбрана длина, которая определялась отсчетами времени за 2001 – 2017 годы. Прогнозирование на один шаг вперед было выполнено для отсчета времени, сопоставленного 2018 году. Сценарный прогноз значений показателя для разных актуальных длин ВР представлен в таблице 2.

Таблица 2 – Сценарный прогноз значений показателя «Объем промышленного производства» Рязанской области

Прогноз на 2018 г.	Период (актуальная длина ВР)				Реальное значение
	2001 – 2017	2006 – 2017	2010 – 2017	2014 – 2017	
Нечеткая линейная регрессия					
Пессимистический	252,240	242,826	242,845	237,875	260,760
Нейтральный	271,167	262,363	254,167	244,235	
Оптимистический	252,240	272,560	263,790	249,575	
Размах (2018 год)	271,167	29,734	20,945	18,059	
Показатель f	0,566	0,579	0,411	0,192	
Нечеткая квадратичная регрессия					
Пессимистический	214,971	232,962	243,284	247,967	260,760
Нейтральный	254,322	253,979	252,245	255,00	
Оптимистический	266,427	264,247	260,760	264,707	
Размах (2018 год)	51,456	31,287	17,476	16,741	
Показатель f	0,747	0,684	0,389	0,136	

Нечеткая регрессия наглядно описывает неопределенность, присущую данным, представленным короткими ВР и позволяет сформировать сценарный прогноз показателей, с учетом существующих оптимистических и пессимистических тенденций развития социально-экономической сферы региона. Значения полученных сценарных прогнозов при различных исходных актуальных длинах ВР позволяют расширить объем доступной информации о текущем состоянии социально-экономической сферы региона.

Использование предложенной модели позволило на основе сформированного сценарного прогноза выявить направления социально-экономического развития, которые были определены как направления с наиболее высоким потенциалом роста (Инвестиции, Объем производства) и наиболее рискованные, требующие особого внимания (Общий ВРП, Строительство, МСП) (акт внедрения АО «Корпорация развития Рязанской области» от 18.08.2020).

Результаты прогнозирования на один шаг вперед с использованием предлагаемой модели прогнозирования подтверждают целесообразность использования уравнения нечеткой регрессии с асимметричными параметрами для оценки уровня неопределенности поведения ВР и формирования сценарных

прогнозов на один шаг вперед для показателей социально-экономической сферы в целях повышения обоснованности принимаемых управленческих решений и их адекватности состоянию социально-экономической сферы региона.

В третьей главе предложена универсальная многокритериальная *FOU*-модель прогнозирования, позволяющая посредством реализации генетического алгоритма, используемого для поиска оптимальных значений параметров модели, сделать сценарный прогноз на один шаг вперед.

Пусть $d(t)$ ($t = 0, 1, 2 \dots m$) – ВР на основе реальных значений прогнозируемого фактора, а $\Delta d(t)$ ($t = 1, 2 \dots m$) – ВР на основе значений приращений этого фактора: $\Delta d(t) = d(t) - d(t - 1)$. Использование ВР $\Delta d(t)$ позволяет повысить точность разрабатываемой модели прогнозирования.

Пусть универсум X , на котором задан ВР $\Delta d(t)$, определен как

$$X = [D_{min} - D_1, D_{max} + D_2], \quad (8)$$

где D_{min} и D_{max} – минимальное и максимальное значения элементов ВР $\Delta d(t)$ ($t = 1, 2 \dots m$) соответственно ($D_{min} = \min_{t=1,m}(\Delta d(t))$, $D_{max} = \max_{t=1,m}(\Delta d(t))$), а D_1 и D_2 – действительные числа, использование которых позволяет разбить универсум X на n интервалов равной длины: x_1, x_2, \dots, x_n .

Интервальное дискретное нечеткое множество второго типа (ИДНМТ2) \tilde{A} , определенное на универсуме X , может быть представлено как:

$$\tilde{A} = u_{\tilde{A}}(x_1)/x_1 + u_{\tilde{A}}(x_2)/x_2 + \dots + u_{\tilde{A}}(x_n)/x_n, \quad (9)$$

где $u_{\tilde{A}}(x) = \underline{u}_{\tilde{A}}(x), \bar{u}_{\tilde{A}}(x)$; $\underline{u}_{\tilde{A}}(x), \bar{u}_{\tilde{A}}(x)$ – «нижняя» и «верхняя» ФП ИДНМТ2, которые характеризуют «отпечаток неопределенности» *FOU*; $u_{\tilde{A}}(x): X \rightarrow [0, 1]$; $u_{\tilde{A}}(x_r)$ ($r = \overline{1, n}$) – значение степени принадлежности интервала x_r универсума X по «нижней» и «верхней» ФП ИДНМТ2.

Лингвистические термы \tilde{A}_r ($r = \overline{1, n}$) на основе ИДНМТ2 задаются как:

$$\begin{aligned} \tilde{A}_1 &= 1/x_1 + \Lambda/x_2 + 0/x_3 + \dots + 0/x_{n-1} + 0/x_n, \\ \tilde{A}_2 &= \Lambda/x_1 + 1/x_2 + \Lambda/x_3 + 0/x_4 + \dots + 0/x_n, \\ \tilde{A}_3 &= 0/x_1 + \Lambda/x_2 + 1/x_3 + \Lambda/x_4 + \dots + 0/x_n, \\ &\dots\dots\dots \\ \tilde{A}_n &= 0/x_1 + 0/x_2 + \dots + 0/x_{n-2} + \Lambda/x_{n-1} + 1/x_n, \end{aligned} \quad (10)$$

где $\Lambda = \alpha_{lower}, \alpha_{upper}$; α_{lower} и α_{upper} – значения «нижней» ФП $\underline{u}_{\tilde{A}}(x_r)$ и «верхней» ФП $\bar{u}_{\tilde{A}}(x_r)$ на интервале x_r ($r = \overline{1, n}$) соответственно.

Лингвистическому терму \tilde{A}_r ($r = \overline{1, n}$) соответствует «отпечаток неопределенности» FOU_r . Если значение приращения ВР принадлежит интервалу x_r ($r = \overline{2, n-1}$), то соответствующий ему «отпечаток неопределенности» FOU_r имеет вид: $FOU_r = \Lambda/\tilde{A}_{r-1} + 1/\tilde{A}_r + \Lambda/\tilde{A}_{r+1}$ ($\Lambda = \alpha_{lower}, \alpha_{upper}$). Если значение приращения ВР принадлежит интервалу x_1 или x_n , то соответствующий такому интервалу «отпечаток неопределенности» утрачивает самый левый или самый правый член в вышеуказанной формуле для FOU_r .

FOU-модель прогнозирования строится в предположении, что для t -го и $(t + 1)$ -го отсчетов времени ВР определены FOU_j и FOU_l соответственно и, поэтому, может быть сформирована нечеткая логическая зависимость (НЛЗ) 1-го порядка: $FOU_j \rightarrow FOU_l$. НЛЗ k -го порядка формируются аналогичным образом.

Группы нечетких логических зависимостей (ГНЛЗ) определяются посредством объединения НЛЗ с одинаковой левой частью в одну группу. Например, может быть определена ГНЛЗ вида: $FOU_{j_k}, FOU_{j_{(k-1)}}, \dots, FOU_{j_1} \rightarrow FOU_{l_1}, FOU_{l_2}, \dots, FOU_{l_g}$.

«Нижняя» и «верхняя» ФП ИДНМТ2 для правой части группы при вычисляются как: $u_{FOU}(x_r) = \max(u_{FOU_{l_1}}(x_r), u_{FOU_{l_2}}(x_r), \dots, u_{FOU_{l_g}}(x_r))$ ($r = \overline{1, n}$).

При прогнозировании значения элемента ВР для $(t + 1)$ -го отсчета времени определяются соответствующие ему ИДНМТ2 и «отпечаток неопределенности» *FOU*. Прогнозируемое значение $f(t + 1)$ для $(t + 1)$ -го отсчета времени находится как сумма известного значения элемента $d(t)$ ВР для t -го отсчета времени и значения центраида для приращения ВР $y(t + 1)$ для $(t + 1)$ -го отсчета времени.

При разработке *FOU*-модели прогнозирования актуальной задачей является поиск оптимальных значений параметров модели прогнозирования, обеспечивающих максимальную точность прогнозирования для повышения обоснованности управленческих решений, принимаемых на основе полученных прогнозов. В качестве таких параметров в *FOU*-модели прогнозирования выступают: действительные числа D_1 и D_2 , используемые при корректировке границ универсума X , число интервалов разбиения n универсума X , порядок k модели прогнозирования, степени принадлежности α_{lower} и α_{upper} . Применение ГА позволяет значительно сократить время поиска оптимальных значений параметров *FOU*-модели прогнозирования посредством решения соответствующей задачи оптимизации.

В качестве критериев для оценки качества модели прогнозирования используются средняя относительная ошибка прогнозирования *AFER*, а также показатель несовпадения тенденций *Tendency*, которые должны быть минимизированы:

$$AFER = \frac{\sum_{t=k+2}^m |(f(t) - d(t))/d(t)|}{m} \cdot 100\%, \quad (11)$$

$$Tendency = \frac{h}{m - k - 1}, \quad (12)$$

где $f(t)$ и $d(t)$ – предсказанное и реальное значения для t -го отсчета времени; m – число элементов ВР; k – порядок модели прогнозирования; h – число отрицательных произведений $(f(t - 1) - f(t)) \cdot (d(t - 1) - d(t))$ при $t = \overline{k + 2, m}$; $m - k - 1$ – общее число произведений $(f(t - 1) - f(t)) \cdot (d(t - 1) - d(t))$.

Значения параметров *FOU*-модели прогнозирования определяются на основе лучшей хромосомы, полученной при реализации многокритериального ГА с учетом принципов Парето-доминирования. При оценке качества предлагаемой *FOU*-модели прогнозирования применяются обоснованно выбранные показатели качества (11) и (12), позволяющие получать (по результатам реализации многокритериального ГА)

адекватные результаты прогнозирования на шаг вперед для ВР социально-экономической сферы региона.

Таблица 3 – Результаты прогнозирования показателей социально-экономической сферы Рязанской области на один шаг вперед

Показатель	Реальное значение (на 2018 год)	Результаты прогнозирования с использованием AR-модели (лаг=8)		Результаты прогнозирования с использованием лучших ARIMA-моделей			Результаты прогнозирования с использованием FOU-моделей		
		Ошибка, %	Прогнозн. значение (на 2018 год)	Тип модели	Ошибка, %	Прогнозн. значение (на 2018 год)	AFER, %	Ошибка, %	Прогнозн. значение (на 2018 год)
Среднемесячная номинальная начисленная заработная плата в целом по региону, руб.	31966,8	1,119	31609,033	(0, 2, 1)	1,871	31368,84	0,959	0,279	32056,05
Реальные денежные доходы населения, % к предыдущему году	98,5	3,018	101,472	(7, 1, 3)	12,250	110,566	2,454	4,608	93,961
Производство сельского хозяйства, млн. руб.	51586,8	21,623	62741,506	(0, 2, 1)	20,354	62086,555	3,241	17,343	60533,59
Индекс производства продукции сельского хозяйства, % к предыдущему году в сопоставимых ценах	97,2	26,287	71,649	(0, 0, 1)	7,241	104,238	5,632	5,172	102,227
Валовой региональный продукт, в основных ценах соответствующих лет; млн. руб.	360600	9,976	396573,154	(0, 2, 1)	5,135	379116,91	1,989	4,428	376568,6
Индекс физического объема валового регионального продукта, % к предыдущему году в сопоставимых ценах	101,6	4,638	106,312	(0, 1, 1)	1,729	99,843	1,256	2,766	98,789
Объем отгруженной продукции (работ, услуг), млн. руб.	333424,4	106,175	-20588,164	(1, 2, 0)	1,080	337025,37	2,380	1,864	327208,26
Индекс промышленного производства, % к предыдущему году	103,8	1,454	105,309	(8, 1, 1)	12,575	90,748	2,377	11,519	91,843
Ввод в действие жилых домов, тыс. кв. м в общей площади	777,5	5,499	734,743	(8, 2, 1)	14,200	667,096	1,396	13,840	669,891
Индекс потребительских цен, декабрь к декабрю предыдущего года, %	106,0	5,902	99,743	(0, 1, 3)	3,262	109,458	1,677	5,184	100,505
Оборот розничной торговли, млрд. руб.	191,62	1,323	189,084	(8, 2, 0)	1,776	195,023	1,498	1,448	194,395

Анализ результатов прогнозирования с использованием предложенной модели показывает, что использование многокритериального ГА для поиска оптимальных значений параметров FOU-моделей прогнозирования позволяет в ряде случаев повысить на 0,9 – 2,1 процентных пункта точность результатов прогнозирования на 1 шаг вперед при приемлемых временных затратах (в остальных случаях предлагаемая модель показала близкие по точности результаты), что позволяет сделать вывод о повышении качества прогнозной информации для принятия управленческих решений и, как следствие, о повышении обоснованности и адекватности решений, принимаемых на основе формируемых прогнозов.

FOU-модель прогнозирования может быть использована не только для прогнозирования будущих значений элементов ВР, но и для прогнозирования значений различных характеристик этого ВР, например, для прогнозирования значений показателя Херста H или фрактальной размерности D . Предварительный расчет таких характеристик позволяет сделать вывод о прогнозируемости (непрогнозируемости) исходного ВР и, следовательно, о целесообразности применения моделей прогнозирования для получения обоснованных прогнозов.

В четвертой главе описана программная реализация разработанных моделей и алгоритмов. Программный продукт – пакет прикладных программ (ППП) «Intellectual

Time Series Analysis» – предназначен для решения задач анализа и прогнозирования показателей социально-экономической сферы, представленных короткими ВР.

ППП состоит из пяти программных модулей.

Программный модуль «Fuzzy identification of the structure transformations» реализует выявление структурных трансформаций в группе ВР с использованием инструментария нечеткой кластеризации (п. 2.1, глава 2).

Программный модуль «Hierarchical identification of the structure transformations» реализует выявление структурных трансформаций в группе ВР с использованием инструментария иерархической кластеризации (п. 2.2, глава 2).

Программный модуль «Phase portrait» реализует построение фазовых портретов ВР с выявлением аттракторов и оценкой их длины с последующим формированием обучающей последовательности данных актуальной длины, которая может быть использована при работе других программных модулей (п. 2.3, глава 2).

Программный модуль «Fuzzy regression model» реализует разработку уравнения нечеткой регрессии, при этом возможна реализация одновременного поиска адекватного типа уравнения нечеткой регрессии, а также значений его параметров с использованием ГА. Программный модуль «Fuzzy regression model» обеспечивает возможность оценки уровня неопределенности поведения ВР в контексте выбранного уравнения нечеткой регрессии, а также возможность прогнозирования на один шаг вперед с получением пессимистического, нейтрального (классического) и оптимистического прогнозов (п. 2.4, глава 2).

Программный модуль «Fuzzy forecasting model» реализует разработку многокритериальной *FOU*-модели прогнозирования, обеспечивает возможность оценки уровня неопределенности поведения ВР в контексте анализа значений параметров «отпечатка неопределенности» *FOU*, а также возможность прогнозирования на один шаг вперед с получением пессимистического, нейтрального (классического) и оптимистического прогнозов (глава 3).

Использование ППП позволяет автоматизировать процессы выявления структурных трансформаций в группах ВР и прогнозирования значений показателей социально-экономической сферы региона, представленных ВР, при подготовке информации для принятия управленческих решений.

ЗАКЛЮЧЕНИЕ

В работе решена актуальная научно-техническая задача, связанная с развитием методов обработки информации для поддержки принятия управленческих решений в социально-экономической сфере, а именно – с развитием средств прогнозирования и анализа показателей социально-экономической сферы, представленных короткими временными рядами, посредством разработки и исследования новых моделей и алгоритмов с использованием инструментария теории нечетких множеств, генетических алгоритмов, кластерного и фазового анализа. Применение предложенных в работе моделей и алгоритмов позволяет получить дополнительную (уточненную) информацию о состоянии и перспективах развития социально-экономической сферы региона для повышения обоснованности и адекватности принимаемых управленческих решений.

Основные результаты диссертационной работы состоят в следующем.

1 Выполнен анализ применимости инструментария теории нечетких множеств и генетических алгоритмов для разработки регрессионных и многокритериальных моделей прогнозирования коротких временных рядов, а также исследование возможностей инструментария кластерного и фазового анализа в задаче выявления структурных трансформаций в группах временных рядов.

2 Предложена модель прогнозирования и анализа коротких временных рядов на основе регрессионного анализа, отличающаяся использованием уравнения нечеткой регрессии с асимметричными параметрами и генетического алгоритма, позволяющая оценить уровень неопределенности поведения временного ряда в контексте выбранного уравнения нечеткой регрессии, а также сформировать на его основе сценарный прогноз показателей социально-экономической сферы региона на один шаг вперед.

3 Предложена универсальная многокритериальная *FOU*-модель прогнозирования, отличающаяся использованием двух минимизируемых критериев, позволяющая посредством реализации генетического алгоритма сделать сценарный прогноз на один шаг вперед и охватывающая частный случай, предполагающий представление временных рядов на основе дискретных нечетких множеств первого типа.

4 Предложена методика оценки вариативности коротких временных рядов социально-экономической сферы, отличающаяся выполнением двухэтапного анализа временного ряда с учетом значений элементов как исходного временного ряда, так и временного ряда, содержащего значения показателя Херста или фрактальной размерности, с применением универсальной многокритериальной *FOU*-модели прогнозирования и позволяющая повысить обоснованность оценки вариативности временного ряда.

5 Разработан алгоритм выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы, отличающийся использованием нечетких *s*-средних и позволяющий, в случае хорошей отделимости кластеров, выполнить анализ поведения групп показателей социально-экономического развития и выявить структурные трансформации (изменения в их динамике) на ранних стадиях появления.

6 Разработаны алгоритмы выявления структурных трансформаций в группах коротких временных рядов социально-экономической сферы, отличающиеся использованием иерархической агломеративной кластеризации, не накладывающей на группу временных рядов жесткие требования хорошей отделимости кластеров, и позволяющие выполнить анализ поведения групп показателей социально-экономического развития и выявить структурные трансформации на ранних стадиях появления в случаях, когда осуществляется анализ показателей социально-экономического развития, относящихся к разным подсистемам социально-экономической системы.

7 Разработан пакет прикладных программ (ППП), реализующий анализ и прогнозирование коротких ВР с применением предложенных алгоритмов и моделей, подтверждающий их адекватность и эффективность.

Результаты прогнозирования и анализа показателей социально-экономической сферы, представленных короткими временными рядами, с применением разработанных моделей и алгоритмов направлены на повышение адекватности и обоснованности принимаемых управленческих решений на основе получаемой информации при приемлемых временных затратах.

Полученные в данной работе результаты показали, что на принимаемые управленческие решения положительно влияют такие факторы, как расширение исходной информации, выявление на ранних этапах структурных трансформаций ВР в регионе, что дает возможность идентифицировать и предотвращать развитие кризисных явлений, повышение точности формируемых прогнозов показателей социально-экономической сферы.

ОСНОВНЫЕ ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в изданиях ВАК:

1. Степанов, М.А. Подход к решению задачи выявления структурных трансформаций в группах временных рядов / Л.А. Демидова, М.А. Степанов // Cloud of Science, 2019. – Т. 6 – № 2. – С. 201–226.
2. Степанов, М.А. Методика выявления структурных трансформаций временных рядов с использованием принципов нечеткой кластеризации / М.А. Степанов // Вестник Рязанского государственного радиотехнического университета, 2019. – № 69. – С. 149 – 159.
3. Степанов, М.А. Многокритериальная оптимизация в моделях прогнозирования на основе интервальных дискретных нечетких множеств второго типа / Л.А. Демидова, М.А. Степанов // Cloud of Science, 2019. Т. 6 – № 3. – С. 401–414.
4. Степанов, М.А. Подход к выбору модели нечеткой регрессии в задаче анализа показателей социально-экономической сферы, представленных временными рядами / Л.А. Демидова, М.А. Степанов // Вестник Рязанского государственного радиотехнического университета, 2020. – № 71. – С. 83 – 97.

Публикации в изданиях, индексируемых в международных базах данных:

5. Stepanov, M.A. Algorithms of Identifying of the Structural Transformations in Time Series Groups / L.A.Demidova, M.A.Stepanov // 8th Mediterranean Conference on Embedded Computing, MECO 2019, 2019. – С. 234 – 237.
6. Stepanov, M.A. Time Series and Its Characteristics Variability Analysis Using the Forecasting Model Based on the Basis of Interval Discrete Fuzzy Sets of the Second Type / L.A Demidova, M.A. Stepanov // 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA), 2019. – P. 241 – 246. DOI: 10.1109/SUMMA48161.2019.8947584.
7. Stepanov, M.A. Identification of the Structural Transformations of Time Series Using Fuzzy Clustering Algorithms / M.A.Stepanov // 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency (SUMMA), 2019. – P. 373 – 376. DOI: 10.1109/SUMMA48161.2019.8947596.
8. Stepanov, M.A. Assessment of the behavior uncertainty level of the short time series / L.A. Demidova, M.A. Stepanov // 2020 Journal of Physics: Conference Series Vol. 1479, Applied Mathematics, Computational Science and Mechanics: Current Problems, 11-13 November 2019, Voronezh, Russian Federation, 2020. – P. 012012. DOI:10.1088/1742-6596/1479/1/012012.
9. Stepanov, M.A. Approach to the Analysis of the Multidimensional Time Series Based on the UMAP Algorithm in the Problems of the Complex Systems Proactive Maintenance / L.A.Demidova, M.A.Stepanov // 2020 International Conference on Information Technologies (InfoTech), Varna, Bulgaria, 2020, pp. 1-4, doi: 10.1109/InfoTech49733.2020.9211008.
10. Stepanov, M.A. Approach to Choosing the Type and Parameters Values of the Fuzzy Regression Model for Time Series / M.A.Stepanov , S.I. Babaev, N.I. Khizrieva// 9th Mediterranean Conference on Embedded Computing, MECO 2020, 2020. DOI: 10.1109/MECO49872.2020.9134318
11. Stepanov, M.A. Multiobjective Forecasting Model Based on the Interval Discrete Type-2 Fuzzy Sets and Genetic Algorithm / L.A.Demidova, M.A.Stepanov // 2020 International Conference on Information Technologies (InfoTech), Varna, Bulgaria, 2020, pp. 1-4, doi: 10.1109/InfoTech49733.2020.9211063.
12. Stepanov, M.A. Data Analysis Using the Nonlinear Dimension Reduction Algorithms // 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2020 DOI: 10.1109/SUMMA50634.2020.9280727.

Публикации в других изданиях:

13. Степанов, М.А. Подход к прогнозированию структурных трансформаций временных рядов / Л.А. Демидова, М.А. Степанов // II Международный научно-технический форум СТНО-2019. Сборник трудов, 2019. С. 75-79.

14. Степанов, М.А. Аспекты расширения набора характеристик в решении задачи кластеризации временных рядов // Новые информационные технологии в научных исследованиях: Материалы XXIV Всероссийской НТК. (НИТ-2019). Рязань, 2019. – С 45–46.

15. Степанов, М.А. Оценка уровня неопределенности поведения короткого временного ряда / Л.А. Демидова, М.А. Степанов // Актуальные проблемы прикладной математики, информатики и механики: сборник трудов МНК, Воронеж: Издательство «Научно-исследовательские публикации», 2020. – С. 1825-1833.

16. Степанов, М.А. Аспекты интеллектуального анализа данных в задаче кластеризации временных рядов / Л.А. Демидова, М.А. Степанов // Новые информационные технологии в научных исследованиях: материалы XXV Юбилейной Всероссийской НТК (НИТ-2020). Рязань, 2020. С. 24-25.

Свидетельства о государственной регистрации программ для ЭВМ:

17. Степанов, М.А. Программный модуль для выявления структурных трансформаций в группах временных рядов (TimesSeriesTransformation) / Л.А. Демидова, М.А. Степанов // Свидетельство о государственной регистрации программы для ЭВМ № 2019660333 (регистрация в Реестре – 05.08.2019).

18. Степанов, М.А. Многокритериальная модель прогнозирования коротких временных рядов на основе интервальных нечетких множеств второго типа (MultiobjectiveTimesSeriesType2FuzzySetsForecasting) / Л.А. Демидова, М.А. Степанов // Свидетельство о государственной регистрации программы для ЭВМ № 2019667588 (регистрация в Реестре – 25.12.2019).

Степанов Максим Анатольевич

**МОДЕЛИ ПРОГНОЗИРОВАНИЯ И АЛГОРИТМЫ АНАЛИЗА
ВРЕМЕННЫХ РЯДОВ ДЛЯ ПОДДЕРЖКИ
ПРИНЯТИЯ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ
В СОЦИАЛЬНО-ЭКОНОМИЧЕСКОЙ СФЕРЕ**

Специальность 05.13.10 – «Управление в социальных и экономических системах»

Автореферат

диссертации на соискание ученой степени
кандидата технических наук

Подписано в печать _____ Формат 60x84¹/16

Усл.печ.л. 1,25. Тираж 100 экз.

Рязанский государственный радиотехнический университет имени В.Ф. Уткина
390005, г. Рязань, ул. Гагарина, д.59/1