

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение
высшего образования
**«УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»**

На правах рукописи



Тимина Ирина Александровна

**АВТОМАТИЗИРОВАННОЕ ПРОЕКТИРОВАНИЕ НА
ОСНОВЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ
НЕЧЕТКИХ ТЕНДЕНЦИЙ МЕТРИК ПРОЕКТНОЙ
ДЕЯТЕЛЬНОСТИ**

Специальность: 05.13.12 – Системы автоматизации
проектирования (промышленность)
05.13.18 – Математическое моделирование,
численные методы и комплексы программ

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель – доктор технических наук, профессор,
Ярушкина Надежда Глебовна

Ульяновск – 2016

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	5
1 ОБЗОР МЕТОДОВ И СРЕДСТВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА	
МЕТРИК ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ	14
1.1 Основные понятия и процессы управления проектной деятельностью	14
1.1.1 Основные понятия и участники управления проектной деятельностью	14
1.1.2 Основные процессы управления проектной деятельностью	16
1.2 Сервисы управления проектами	24
1.3 Обзор средств организации проектной информации. Понятие системы контроля версий	30
1.4 Непрерывная интеграция.....	32
1.5 Описание автоматизированной системы отдела технической документации на примере крупной проектной организации ФНПЦ АО «НПО Марс»	34
1.6 Основные методы интеллектуального анализа временных рядов	41
1.6.1 Базовые понятия анализа временных рядов	41
1.6.2 Методы интеллектуального анализа временных рядов.....	43
1.7 Нечеткие модели временного ряда.....	46
1.7.1 Основы нечеткого моделирования временных рядов.....	46
1.7.2 Обзор методов нечеткого моделирования временных рядов.....	48
1.7.3 Нечеткая тенденция. Виды НТ (элементарная, локальная, общая).....	50
1.8 Мера ассоциации локальных трендов ВР - угловой коэффициент скользящей аппроксимации	56
1.9 Оценки качества ВР	59
1.9.1 Абсолютные отклонения прогнозов	59
1.9.2 Процентные (относительные) отклонения прогнозирования	61
1.10 Постановка задачи исследования	63
2 МОДЕЛЬ И АНАЛИЗ МЕТРИК ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ	64
2.1 Модель проектной деятельности на основе метрик систем контроля версий	64
2.2 Построение коэффициента сходства ВР. Выявление степени зависимости между ВР метрик проектной деятельности	66
2.3 Кластеризация временного ряда.....	68
2.3.1 FCM-алгоритм кластеризации.....	68
2.3.2 Кластеризация ВР по доминирующей тенденции.....	71

2.4 Прогнозирование ВР метрик проектной деятельности	72
2.4.1 Прогнозирования ВР в терминах нечетких тенденций	74
2.4.2 Применение F-преобразования в задаче определения сходных временных рядов.....	75
2.4.3 Метод и алгоритм прогнозирования ВР на основе нечетких тенденций.....	78
2.5 Алгоритм коррекции гипотез прогноза зависимых временных рядов	80
2.5.1 Прогнозирование временного ряда посредством алгоритма адаптации	80
2.5.2 Правила анализа зависимости между ВР	83
2.6 Коэффициент подобия как способ оценки влияния наличия нечетких тенденций шаблона на прогноз	83
2.7 Выводы.....	85
 3 РЕАЛИЗАЦИЯ ПРОГРАММНОЙ СИСТЕМЫ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ СОБЫТИЙ СКВ	86
3.1 Информационная структура программной системы анализа и прогнозирования метрик (ПС АПМ).....	86
3.1.1 Моделирование процесса анализа метрик проекта в нотации IDEF0/IDEF3 ..	86
3.1.2 Структурно-функциональная схема интеллектуального хранилища.....	88
3.2 Обоснование инструмента реализации	90
3.3 Структуры данных ПС АПМ	91
3.3.1 Структура входных данных.....	91
3.3.2 Структура выходных данных	92
3.4 Иерархия классов ПС АПМ	92
3.5 Моделирование ПС АПМ.....	97
3.5.1 Диаграмма вариантов использования.....	97
3.5.2 Диаграмма состояний.....	99
3.5.3 Диаграмма последовательности.....	101
3.5.4 Диаграмма компонентов.....	104
3.6 Выводы.....	105
 4 ОЦЕНКА АДЕКВАТНОСТИ МОДЕЛЕЙ НА БАЗЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ ЭФФЕКТИВНОСТИ СРЕДСТВА АВТОМАТИЗАЦИИ ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ	106
4.1 Эффекты внедрения программной системы анализа и прогнозирования метрик проектной деятельности	106
4.2 План экспериментов	107

4.3 Исследование временных рядов метрик на подобие и корреляцию	107
4.3.1 Применение F-преобразования	111
4.3.2 Сравнение методов преобразования с методами углового коэффициента сходства.....	115
4.4 Прогнозирование временных рядов на основе сходства поведения. Показатели критериев качества прогноза	116
4.4.1 «Измерение» метрик программных проектов с открытым кодом.....	127
4.5 Эффективность ПСАП СКВ.....	136
4.6 Выводы.....	146
ЗАКЛЮЧЕНИЕ	147
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ	148
СПИСОК ЛИТЕРАТУРЫ.....	149
ПРИЛОЖЕНИЕ А. АКТ ВНЕДРЕНИЯ ПСАП СКВ В ФНПЦ АО «НПО «МАРС» (Г. УЛЬЯНОВСК)	159
ПРИЛОЖЕНИЕ В. АКТ СДАЧИ-ПРИЕМКИ НА ВЫПОЛНЕНИЕ РАБОТЫ «РАЗРАБОТКА НЕЧЕТКИХ МОДЕЛЕЙ И РЕАЛИЗАЦИЯ НЕЧЕТКИХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ» ООО «ЭВЕРЕСТ РЕСЕРЧ» (Г. МОСКВА)	161
ПРИЛОЖЕНИЕ С. СВИДЕТЕЛЬСТВО О ГОСУДАРСТВЕННОЙ РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ	163
ПРИЛОЖЕНИЕ D. СЕРТИФИКАТ ПОБЕДИТЕЛЯ МЕЖДУНАРОДНОГО КОНКУРСА ПО ВРЕМЕННЫМ РЯДАМ «ВЫЧИСЛИТЕЛЬНЫЙ ИНТЕЛЛЕКТ В ПРОГНОЗИРОВАНИИ».....	165
ПРИЛОЖЕНИЕ Е. ДАННЫЕ ПОСТУПЛЕНИЯ ОТЧЕТОВ ПО ИЗДЕЛИЯМ ФНПЦ АО «НПО «МАРС»	166
ПРИЛОЖЕНИЕ F. ДАННЫЕ МЕТРИК ПРОЕКТА «РАЗРАБОТКА НЕЧЕТКИХ МОДЕЛЕЙ И РЕАЛИЗАЦИЯ НЕЧЕТКИХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ»	173
ПРИЛОЖЕНИЕ G. РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ	175

ВВЕДЕНИЕ

У большинства крупных проектных предприятий имеются значительные архивы различного рода проектов. Наряду с электронными архивами конструкторско-технологической информации при разработке сложных программно-аппаратных комплексов для хранения проектных документов и программного кода широко используется система контроля версий (СКВ). В СКВ регистрируется множество различных событий, которые касаются сущностей проекта. Некоторые проекты используют ранее разработанные решения, так как это значительно сокращает сроки выполнения. Проекты часто бывают взаимосвязаны между собой. Жизненный цикл сложной системы – продолжительный и включает в себя несколько модернизаций, каждую из которой можно рассматривать как проект. В связи с этим в крупной проектной организации массив действующих активных проектов очень большой. Естественным агрегатором сущностей всех проектов является СКВ, так как она вбирает в себя всю информацию по работе над проектом. Поэтому в крупных проектных организациях возникает многоуровневая модель управления. Возникает сложный контур метауправления, который управляет совокупностью проектов на протяжении длительного времени. Появляется проблема мониторинга состояния проектной деятельности для главного инженера (конструктора) сложных проектов. Поэтому важно иметь инструмент измерения характеристик проектной деятельности, который позволяет отследить динамику совокупности проектов, идентифицировать проблемы и своевременно принять корректирующие регулирующие воздействия, в том числе перераспределение ресурсов между проектами.

Формой метрической характеристики проекта может являться временной ряд, который будет получаться на основе выделенных сущностей за определенный период времени по каждому проекту.

В процессе реализации проекта от момента утверждения стабильной версии до выпуска нового релиза ставится определенное количество задач по введению

новых функций системы, исправлению возникающих при разработке ошибок в связи с реализацией новых функций и добавлению улучшений (исправлений ошибок). Введение новых функций оказывает влияние на добавление улучшений. Следовательно, метрика «улучшения» связана с метрикой «новые функции». Аналогичная взаимосвязь существует между появлением ошибок и улучшениями. Метрика «ошибки» является предиктором для метрики «добавленные улучшения». В зависимости от конкретного проекта перед проектировщиком могут стоять различные задачи – это либо выявление и исправление ошибок, либо развитие проекта за счет усовершенствования.

Одной из задач управления проектом является задача прогнозирования. Прогноз с учетом основных задач ведения проекта (например, прогноз получения ошибок) покажет успешность проекта и позволит руководителю верно распределять задания сотрудникам.

Для решения задачи прогнозирования, в виду большого объема хранимой информации, требуется обширный анализ метрик проектных процессов и документов, который сможет реализовать поиск схожих сущностей проектных документов, выявить степень их подобия и на основе этих показателей получить будущее значение. Таким образом, появляется задача создания средства для управления эффективности проектом, которое автоматизирует процессы классификации, схожести всех имеющихся сущностей проектов предприятия для последующего прогноза значений.

В настоящий момент времени отсутствуют методы прогнозирования метрик сущностей СКВ на основе моделей проектной деятельности. Современные СКВ не ориентированы на анализ и прогнозирование метрик проекта, содержащего неоднородные, нелинейные и разнонаправленные данные и связи. Поэтому возникает проблема формирования автоматизированного метода прогнозирования метрик системы управления версиями проектов, в том числе на основе анализа зависимости (корреляции) метрик. Таким образом, для реализации прогнозирования поступления сущностей проекта в интеллектуальном проектном репозитории САПР требуется содержательный анализ исследуемых метрик.

Современный проектный репозиторий, включающий СКВ, необходимо интерпретировать как интеллектуальное хранилище проекта.

Цель работы

Целью работы является построение эффективного средства управления совокупностью проектов крупной проектной организации на основе анализа и моделирования состояния метрик проектов программно-аппаратных комплексов.

Объектом и материалом исследования послужили статистика событий системы контроля версий различных проектов, статистика деятельности отдела технической документации ФНПЦ АО «НПО «Марс» (Приложение Е), разработка программной системы ООО «Эверест Ресерч», проекты с открытым кодом (Приложение F).

Задачи исследования

1. Выполнить сравнительный анализ методов и средств построения различных видов систем контроля версий, средств управления проектами.
2. Выполнить научный обзор методов анализа временных рядов.
3. Разработать формализованную модель состояния проектной деятельности на основе метрик проектов, загруженных в систему контроля версий проектного репозитория.
4. Разработать средство анализа состояния метрик сущностей СКВ проектов программного обеспечения как инструмент управления работами организации.
5. Разработать нечеткие модели и алгоритмы прогнозирования временных рядов метрических характеристик проекта.
6. Разработать модели и алгоритмы прогнозирования временных рядов метрик проекта на основании мер сходства.
7. Разработать комплекс программ моделирования как подсистему автоматизации СКВ.

Методы исследования: теория нечетких систем и мягких вычислений; методы анализа временных рядов, корреляционного анализа; теория кластеризации; объектно-ориентированный подход при создании комплекса программ.

Научная новизна

1. Средства анализа процессов реализации проектных работ на основе обработки временных рядов индикаторов состояния метрик сущностей СКВ проектов программного обеспечения крупной проектной организации, отличающиеся встроенными возможностями обработки совокупности временных рядов, являются новыми.
2. Предложенная структура автоматизации организации метрик сущностей СКВ, отличающаяся наличием компоненты анализа временных рядов индикаторов состояния СКВ, является новой.
3. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечеткого моделирования, отличающиеся использованием системы нечетких правил для различных гипотез прогноза, являются новыми (предложены три гипотезы прогноза: гипотеза сохранения тенденции, гипотеза смены тенденции и гипотезы смены тенденции на заданный период).
4. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечетких мер сходства временных рядов, отличающиеся предложенными индикаторами сходства и процедурами их построения и внесения поправок в прогноз, являются новыми.

Практическая значимость

Разработанная программная система анализа динамики метрик проектной деятельности используется в работе отдела технической документации ФНПЦ АО «НПО «Марс» (Приложение А), в практике работы ООО «Эверест Ресерч» (г. Москва) (Приложение В) и позволяет оперативно управлять совокупностью текущих проектов.

Основания для выполнения работы

Результаты диссертационной работы использовались в ряде НИОКР, выполненных в Ульяновском государственном техническом университете, направленных на решение научно-технических задач. К наиболее важным результатам следует отнести:

1. Участие в выполнении гранта РФФИ №14-01-31092 «Разработка метода интеллектуального анализа многомерных нечетких временных рядов» (руководитель).

2. Участие в выполнении гранта РФФИ №13-01-00324 «Исследование формальных методов грануляции слабоструктурированных информационных ресурсов на основе онтологии предметной области».

3. Участие в выполнении гранта РФФИ №14-07-00247 «Методология моделирования и прогнозирования локальных тенденций временных рядов в задачах анализа больших данных».

4. Участие в выполнении гранта РФФИ №15-41-02413 «Интеллектуальный анализ временных рядов на основе нечетких онтологий, извлекаемых из Интернет-ресурсов».

5. Участие в реализации государственного задания №2014/232 на выполнение государственных работ в сфере научной деятельности Минобрнауки России по проекту «Разработка нового подхода к интеллектуальному анализу слабоструктурированных информационных ресурсов».

Достоверность результатов диссертационной работы

Достоверность научных положений, выводов и рекомендаций подтверждена результатами вычислительных экспериментов, а также результатами использования созданных алгоритмов и методик в повседневной работе отдела технической документации ФНПЦ АО «НПО «Марс», в практике работы ООО «Эверест Ресерч» (г. Москва).

Основные положения, выносимые на защиту

1. Средства анализа событий процессов реализации проектных работ на основе обработки временных рядов метрик сущностей системы контроля версий крупной проектной организации являются эффективным инструментом управления совокупностью проектов.

2. Система анализа временных рядов метрик сущностей системы контроля версий является необходимой компонентой автоматизации современных проектных репозиторий.

3. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечеткого моделирования является эффективными.

4. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечетких мер сходства временных рядов является эффективными.

Апробация результатов исследования

Основные положения и результаты диссертации представлены в журналах из Перечня, рекомендованного ВАК РФ: «Наукоемкие технологии», «Автоматизация процессов управления», «Радиотехника»; докладывались, обсуждались и получили одобрение: на международной вузовской конференции Interactive systems and technologies: the problem of Human-Computer Interaction (Collection of scientific papers) 2011, 4-й Всероссийской научно-технической конференции аспирантов, студентов и молодых ученых ИВТ-2012, Тринадцатой национальной конференции по искусственному интеллекту с международным участием КИИ-2012, IV Всероссийской школы-семинара аспирантов, студентов и молодых ученых ИМАП-2012, 5-й Всероссийской научно-технической конференции аспирантов, студентов и молодых ученых ИВТ-2013, VII-ой Международной научно-практической конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте» 2013, 48й Научно-технической конференции «Вузовская наука в современных условиях», УлГТУ-2014, SYRCoSE 2014, The 8th Spring / Summer Young Researchers' Colloquium on Software Engineering, VI-ой Всероссийской научно-практической конференции Нечеткие системы и мягкие вычисления–2014 (НСМВ–2014), Четырнадцатой национальной конференции по искусственному интеллекту с международным участием (КИИ-2014), 2-ой Международный Поспеловский симпозиум «Гибридные и синергетические интеллектуальные системы ГИСИС'2014», 1-ой всероссийской научно-практической конференции "Прикладные информационные системы", УлГТУ-2015, 49й Научно-технической конференции «Вузовская наука в современных условиях», УлГТУ-2015, VIII-ой Международной научно-практической конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте» (18-20 мая 2015 г., Коломна); отмечены сертификатом «Combination of

Fuzzy and Exponential Models», подтверждающим занятое первое место в международных соревнованиях «The International Time Series Competition “Computational Intelligence in Forecasting”» (в соавторстве) (Приложение D); представлены двумя свидетельствами о государственной регистрации программного обеспечения №2014612213, №2015617046 (Приложение С). Апробация результатов диссертации осуществлялась в рамках проекта РФФИ № 14-01-31092 «Разработка метода интеллектуального анализа многомерных нечетких временных рядов». Основные положения и результаты неоднократно докладывались и обсуждались на научно-технических конференциях УлГТУ «Вузовская наука в современных условиях».

По результатам работы было опубликовано 26 статей, в том числе, 7 в журналах из Перечня, рекомендованного ВАК РФ. Получены 2 свидетельства о государственной регистрации программного обеспечения (Приложение С).

Структура и объем диссертационной работы

Диссертационная работа состоит из 176 страниц машинописного текста, содержащих введение, четыре главы и заключение, 41 таблицы, 33 рисунка, список литературы из 145 наименований, 7 приложений.

Во **введении** рассмотрена актуальность создания автоматизированного метода прогнозирования состояния событий системы управления версиями проекта на основе схожести сущностей проектов, формируются цель и задачи работы, определяются теоретическая значимость и практическая ценность исследования, а также структура работы, сформулированы положения, выносимые на защиту.

В **первой главе** сделан сравнительный анализ существующих методов и средств построения различных видов систем контроля версий. Выполнен сравнительный анализ средств управления проектами: рассмотрены основные понятия, структура и участники проектной деятельности, принципы и классификация проектов, группы процессов управления проектом, процессы инициации проектов, процессы планирования проектов, процессы исполнения,

процессы мониторинга и управления изменениями, процесс завершения проектов, общая структура систем управления версиями проектов.

Сделаны выводы, что существующие методы и средства построения систем контроля версий не включают анализ метрик состояния разработки проекта, необходимых для управления проектными работами, а так же отличаются отсутствием возможности обработки совокупности проектов.

Выполнен научный обзор методов анализа временных рядов метрик проектов. Приведено краткое описание основных задач интеллектуального анализа, алгоритмов преобразования нечетких тенденций, кластеризации, аргументирована возможность использования нечеткого моделирования и нечетких методов при прогнозировании состояния событий системы контроля версий.

Во **второй главе** предлагается новая архитектура системы управления версиями проекта, включающая компоненту анализа временных рядов индикаторов состояния разработки проектных решений в качестве инструмента управления проектом. Разработан метод и алгоритм прогнозирования тенденций временных рядов на основе нечеткого моделирования, который характеризуется использованием системы нечетких правил для различных гипотез прогноза.

Для обработки совокупности сущностей проекта разработаны модель и алгоритмы прогнозирования тенденций временных рядов на основе нечетких мер их сходства за счет корректировки прогноза состояния событий системы управления версиями проекта.

В **третьей главе** изложено описание структурно-функционального решения программной системы, осуществляющей моделирование и прогнозирование временного ряда метрик проекта с учетом взаимного влияния метрик предиктора.

В **четвертой главе** представлено применение методики корректировки прогнозирования временных рядов индикаторов состояния разработки проекта в системе контроля версий. В качестве примера был рассмотрен проект MongoDB, а именно главная его программная часть Core Server. Данными по событиям проекта

были взяты с ресурса jira.mongodb.org. Представлены результаты внедрения средства автоматизации проектной деятельности в АО ФНПЦ «НПО МАРС».

В **заключении** приведены основные результаты исследований, представленные в диссертационной работе.

В **приложения** входят:

- А. Акт внедрения ПСАП СКВ в ФНПЦ АО «НПО «Марс» (г. Ульяновск).
- В. Акт сдачи-приемки на выполнение работы «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов» ООО «Эверест Ресерч» (г. Москва).
- С. Свидетельство о государственной регистрации программы для ЭВМ.
- Д. Сертификат победителя международного конкурса по временным рядам «Вычислительный интеллект в прогнозировании».
- Е. Данные поступления отчетов по изделиям ФНПЦ АО «НПО «Марс».
- Ф. Данные метрик проекта «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов».
- Г. Результаты вычислительных экспериментов.

1 ОБЗОР МЕТОДОВ И СРЕДСТВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА МЕТРИК ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ

1.1 Основные понятия и процессы управления проектной деятельностью

1.1.1 Основные понятия и участники управления проектной деятельностью

Проект – ограниченное во времени мероприятие, предназначенное для создания уникальных продуктов.

Продукт проекта – объект, появляющийся в результате исполнения проекта.

Участники проекта – лица или организации, либо активно участвующие в проекте, либо те, на чьи интересы могут повлиять результаты исполнения или завершения проекта. Участники также могут влиять на цели и результаты проекта.

Управление проектами – приложение знаний, навыков, инструментов и методов к операциям проекта для удовлетворения требований, предъявляемых к проекту.

План-график проекта – документ, утвержденный Генеральным директором, предназначенный для исполнения и управления проектом.

Проектный документ – материальный объект, который фиксирует или подтверждает какую-либо информацию и может быть включен в определенное собрание [89].

Постановка задачи проекта – проектный документ (например, техническое задание) или комплект документов, описывающий постановку задачи на разработку Продукта проекта.

Проектный офис – специально образованное подразделение для организации, координации и руководства проектной деятельностью.

Документовед Проектного офиса – сотрудник, отвечающий за прием, выдачу и хранение всей документации, полученной в ходе проектной деятельности, а также за организацию документооборота Проектного офиса.

Оператор информационной безопасности – организация, ответственная за формирование и реализацию единой технологической политики построения инфраструктуры электронного документооборота предприятия.

Заинтересованные стороны проекта – это лица или организации (например, заказчики, исполнитель), которые активно участвуют в проекте или интересы которых могут быть затронуты как положительно, так и отрицательно в ходе исполнения или в результате завершения проекта.

Владелец процесса – исполнитель, несущий полную ответственность за качество выполнения процесса.

Заказчик - подразделение или должностное лицо, являющееся основным потребителем результатов продукта проекта.

Функциональный Заказчик – подразделение или должностное лицо, являющееся потребителем продукта проекта и предъявляющее к продукту проекта функциональные требования.

Генеральный директор – лицо, определяющее стратегическое развитие проектов предприятия.

Главный конструктор – специально назначенный сотрудник, наделенный необходимыми полномочиями и выполняющий все функции и по организации исполнения проекта.

Пул главных конструкторов – штатные сотрудники отдела проектного управления.

Устав проекта – документ утвержденный Главным конструктором, с описанием первоначальных требований, удовлетворяющих потребностям и ожиданиям заинтересованных сторон проекта.

Исполнитель – организация или подразделение, осуществляющее исполнение проекта.

Прямой руководитель – должностное лицо, которому подчиняется сотрудник или подразделение согласно организационно-штатной структуре. Ближайший прямой руководитель называется непосредственным руководителем.

Рабочая группа проекта – временно создающаяся организационная единица, являющаяся координационным центром управления отдельным проектом.

Эксперт проекта – профессионал, обладающий компетенцией в области, относящейся к проекту.

Эксплуатирующая (Сопровождающая) организация – организация, осуществляющая техническую эксплуатацию и функциональное сопровождение Продукта проекта после его ввода в промышленную эксплуатацию.

Руководитель направления деятельности организации – лицо, наделенное полномочиями по решению вопросов, выходящих за рамки полномочий Главного конструктора и Рабочей группы проекта, и представляющее интересы проекта перед Генеральным директором.

1.1.2 Основные процессы управления проектной деятельностью

Для представления всех этапов прохождения проекта в проектной организации, рассмотрим структуру и персонал, участвующий в реализации проектов.

Проектное управление предполагает использование трехуровневой модели управления представленной на рисунке 1.1 [88]:

Стратегический уровень

Генеральный
директор

Оперативный уровень

Проектный офис

Тактический уровень

Рабочая группа проекта



Рисунок 1.1 – Схема управления проектами

На рисунке 1.2 представлена общая схема взаимодействия между участниками проектной деятельности:

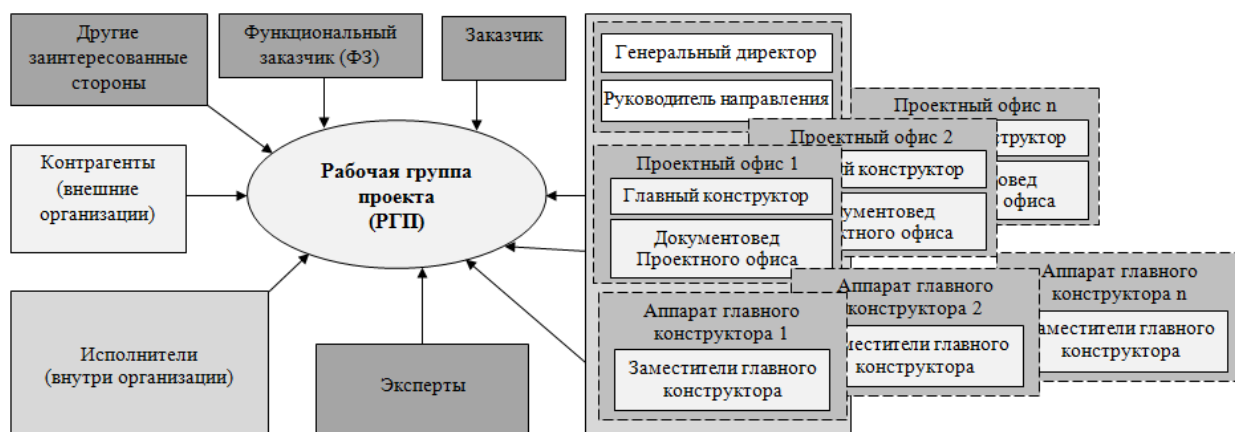


Рисунок 1.2 – Схема взаимодействия между участниками проектной деятельности

Генеральный директор отвечает за разработку и реализацию стратегии проектной деятельности, а также отвечает за реализацию всех проектов перед советом директоров. Генеральный директор наделен следующими функциями:

1. Формирует стратегию проектной деятельности и обеспечивает ее выполнение;
2. Контролирует показатели работы Проектного офиса;

3. Принимает окончательные решения по всем вопросам, связанным с проектной деятельностью;
4. Принимает решения о дополнительном анализе заявок на проекты и решения об инициации проектов;
5. Утверждает Устав проекта и План-график проекта;
6. Назначает Руководителей направлений;
7. Является окончательным арбитром в разрешении всех споров связанных с проектной деятельностью.

Руководитель направления деятельности организации утверждается Генеральным директором. Руководитель направления несет ответственность перед Генеральным директором за достижение целей проекта. Функции Руководителя направления:

1. Совместно с Руководителем Проектного офиса назначает ГИП (главного инженера проекта), главного конструктора;
2. Иницирует смену ГИП;
3. Анализирует отчетность по проекту;
4. Принимает решения по вопросам, связанным с управлением проектом, выходящим за рамки компетенции ГИП;
5. Иницирует процедуру изменения План-графика проекта;
6. Согласует изменения содержания проекта;
7. Согласует изменение целей проекта;
8. Иницирует процедуру приостановления проекта;

В рамках проектной деятельности **Проектный офис** подчиняется Генеральному директору. **Состав Проектного офиса:**

1. Руководитель Проектного офиса;
2. Документовед Проектного офиса;
3. Главные конструктора.

Проектный офис осуществляет техническую и методологическую поддержку проектной деятельности, а также контроль над исполнением проектов. Функции Проектного офиса:

1. Определение и разработка методологии управления проектами;
2. Управление принципами, процедурами, шаблонами проектов и другой общей документацией;
3. Анализ исполнения проектов и соответствия процессов управления проектами стандарта проектной организации;
4. Централизованное управление для всех проектов Проектного офиса;
5. Формирование и рассылка отчетности по проектной деятельности;
6. Ведение архива проектов – накопление, хранение и предоставление согласно утвержденному регламенту документации по проектам (Уставы, План-график, паспорта мероприятий).

Руководитель Проектного офиса отвечает за организацию работы Проектного офиса. Руководитель Проектного офиса несет ответственность перед Главным конструктором, Руководителем направления и перед Генеральным директором за результаты деятельности Проектного офиса. Руководитель Проектного офиса принимает оперативные решения по проектной деятельности. Функции руководителя Проектного офиса:

1. Обеспечивает работоспособность Проектного офиса;
2. Контролирует соблюдение общих методологических требований по управлению проектами;
3. Управляет сотрудниками Проектного офиса;
4. Совместно с Руководителем направления деятельности организации назначает Главных конструкторов;
5. Контролирует проектную деятельность Главных конструкторов;
6. Принимает решения по операционным вопросам проектной деятельности;
7. Отвечает за методологическое наполнение проекта.

Документовед Проектного офиса назначается Руководителем Проектного офиса. Документовед Проектного офиса находится в прямом подчинении Руководителя Проектного офиса. Функции Документоведа Проектного офиса:

1. Прием, выдача документов и поддержание Архива в актуальном состоянии;
2. Проверка правильности оформления проектной документации;
3. Подготовка и своевременное предоставление отчетности по проектной деятельности.

Главного конструктора (ГК) назначает Руководитель направления деятельности организации совместно с Руководителем Проектного офиса. ГК несет ответственность перед Руководителем направления, Руководителем Проектного офиса и Заказчиком за результат исполнения проекта. **Главный конструктор** отвечает за:

1. Достижение целей проекта;
2. Удовлетворение требований Заинтересованных сторон проекта в запланированные сроки, в рамках бюджета и с соответствующим качеством;
3. Своевременное предоставление отчетной информации по проекту всем Заинтересованным сторонам проекта в согласованные сроки;
4. Качество проектной документации;
5. Эффективное взаимодействие всех участников проекта.

В рамках руководства проектом **Главный конструктор** находится в подчинении Руководителя направления деятельности организации. В рамках предоставления отчетности по проекту **Главный конструктор** подчиняется Руководителю Проектного офиса. **Функции Главного конструктора:**

1. Уточняет содержание проекта;
2. Осуществляет совместно с Рабочей группой проекта детальное планирование проекта;
3. Формирует План-график проекта;
4. Разрабатывает Устав проекта;
5. Следит за соблюдением требуемого качества результатов проекта, организует необходимые экспертизы в проекте;
6. Организует работу исполнителей;

7. Контролирует сроки выполнения работ по проекту;
8. Участвует в подготовке и отслеживает бюджет проекта;
9. Управляет системой коммуникаций в проекте;
10. Принимает решения по оперативным вопросам в проекте;
11. Иницирует процедуру изменения содержания проекта;
12. Организует документооборот Рабочей группы проекта;
13. Организует сбор, обработку и своевременное предоставление отчетности по проекту;
14. Готовит совещания Рабочей группы проекта;
15. Формулирует технические задания (ТЗ) Исполнителю и участие в контроле их исполнения;
16. Анализирует запросы на изменения на этапах проектирования и выполнения;
17. Разрабатывает модель требований Функционального Заказчика;
18. Контролирует непротиворечивость функциональных требований при участии в проекте нескольких Функциональных Заказчиков.

При создании программно-аппаратных комплексов с внедрением на объектах создается аппарат главного конструктора, который включает



Рисунок 1.3 – Взаимодействие участников управления проектами

заместителя главного конструктора по технологии разработки программного обеспечения (ПО), заместителя главного конструктора *по технологии создания технических средств (ТС)*, по разработке документации, по объектовым работам. Состав аппарата главного конструктора (ГК) приведен на рисунке 1.3. [83]

1. **Сфера ответственности заместителя ГК по технологии разработки ПО:**
 - автоматизация разработки и проектирования ПО;
 - внедрение технологий коллективной разработки;
 - автоматизация тестирования ПО;
 - внедрение CASE-систем;
 - внедрение новых методологий и технологий по жизненному циклу ПО.
2. **Направления работы заместителя ГК по технологии создания технических средств (ТС):**
 - САПР металлоконструкций;
 - Система управления данными об изделии (PDM/PLM);
 - САПР схемной и системной документации;
 - САПР испытаний и расчетов;
 - системы технической подготовки производства;
 - САПР ЧПУ.
3. **Заместитель ГК по разработке документации** отвечает за:
 - состав документации;
 - документирование функциональных требований к проекту;
 - консультирование по вопросам технических требований;
 - сдачу документации в архив и заказчику;
 - своевременную разработку в полном объеме.
4. **Заместитель ГК по объектовым работам** отвечает за:
 - работу на объектах;
 - за доставку изделия на объект;
 - пусконаладочные работы на объекте.

Исполнители по проектам определяются путем конкурсных процедур (в случае если исполнителем является внешняя организация) или по функциональной направленности деятельности подразделений участвующих в проектной деятельности согласно данным методическим рекомендациям. Основной задачей Исполнителя является реализация Продукта проекта. Функции Исполнителя:

1. Формирует План-график выпуска Продукта проекта в соответствии с План-графиком проекта;
2. Предлагает для рассмотрения Функциональному Заказчику проекта обоснованные изменения к Продукту проекта;
3. Исполняет работы необходимые для получения Продукта проекта;
4. Передает Функциональному Заказчику промежуточные и окончательные результаты проекта согласно действующего План-графика.

Заказчик определяет цели и задачи проекта. Функции Заказчика:

1. Определяет цели и задачи проекта;
2. Формулирует технико-экономическое обоснование к Продукту проекта;
3. Согласует требования к продукту проекта (представленные в ТЗ);
4. Контролирует процесс реализации проекта (сроки, качество);
5. Согласовывает промежуточные результаты работ по проекту;
6. Принимает результаты проекта.

Представитель Заказчика определяет назначение, технико-экономическое обоснование и функциональные требования к Продукту проекта. В проекте может присутствовать несколько Функциональных Заказчиков. Задачи Функционального Заказчика:

1. Формирует функциональные требования к Продукту проекта;
2. Согласует требования к продукту проекта (представленные в ТЗ);
3. Принимает результаты проекта;
4. Иницирует изменения целей проекта;
5. Согласовывает промежуточные результаты работ по проекту.

Рабочая группа проекта согласует тактические решения по проекту, выходящие за рамки обязанностей Главного конструктора. Рабочая группа проекта формируется на этапе инициации проекта и является координирующим центром управления отдельным проектом. Совещания аппарата проводятся в форме заседаний, по результатам которых оформляется протокол. Функции Рабочей группы проекта:

- 1) разрешает вынесенные на заседание Рабочей группы проекта проблемы и конфликты проекта;
- 2) рассматривает и утверждает изменения в содержании проекта в рамках своих полномочий;
- 3) проводит приемку промежуточных и окончательных результатов проекта;
- 4) протоколирует заседания Рабочей группы.

Рабочая группа проекта расформировывается по окончании проекта. Рабочая группа проекта имеет следующую структуру:

- 1) руководитель Рабочей группы;
- 2) заместитель руководителя Рабочей группы;
- 3) секретарь Рабочей группы;
- 4) члены Рабочей группы.

Членами Рабочей группы проекта являются уполномоченные представители Проектного офиса, Заказчика, Функционального Заказчика, Исполнителя, Эксплуатирующей организации, другие Заинтересованные стороны проекта. Любой член Рабочей группы может привлекать к работе над проектом экспертов и своих представителей. Эксперт проекта предоставляет независимые экспертные оценки по ходу проекта.

1.2 Сервисы управления проектами

Согласно [120] **управление проектами** — в соответствии с определением национального стандарта *ANSI (American National Standards Institute) PMBoK (Project Management Body of Knowledge)* [104] — область деятельности, в ходе

которой определяются и достигаются четкие цели *проекта* при балансировании между объемом работ, ресурсами (такими как деньги, труд, материалы, энергия, пространство и др.), временем, качеством и рисками. Ключевым фактором успеха проектного управления является наличие чёткого заранее определённого плана, минимизации рисков и отклонений от плана, эффективного управления изменениями.

Для анализа работы над проектом, решения поставленных задач и оценки появляющихся рисков существует программное обеспечение для управления проектами.

Программное обеспечение для управления проектами — комплексное программное обеспечение, которое состоит из приложений для планирования задач, составления расписания, контроля цены и управления бюджетом, распределения ресурсов, совместной работы, общения, быстрого управления, документирования и администрирования системы, которое используются совместно для управления крупными проектами.

Задачи сервиса управления проектами [97]:

1. Одной из наиболее распространенных возможностей является возможность **планирования** событий и управления задачами. Требования могут различаться в зависимости от того, как используется инструмент. Наиболее распространенными являются:

- планирование различных событий, зависящих друг от друга;
- идентификация крупных составных частей проекта (вехи проекта) и их декомпозиция, посредством которой создается структура декомпозиции работ, также называемая иерархической структурой работ;
- планирование расписания работы сотрудников и назначение ресурсов на конкретные задачи;
- расчет времени, необходимого на решение каждой из задач;
- сортировка задач в зависимости от сроков их завершения;
- презентация графика работ по проекту;
- управление несколькими проектами одновременно.

2. Управление данными и предоставление информации

Программное обеспечение для управления проектами предоставляет большое количество требуемой информации, такой как:

- список задач для сотрудников и информацию распределения ресурсов;
- обзор информации о сроках выполнения задач;
- ранние предупреждения о возможных рисках, связанных с проектом;
- информации о рабочей нагрузке;
- информация о ходе проекта, показатели и их прогнозирование.

3. Управление коммуникациями команды проекта

- обсуждение и согласование рабочих вопросов проекта;
- фиксация проблем проекта и запросов на изменения, их обработка;
- ведение рисков проекта и управление ими;
- предоставление доступа к информации о ходе проекта в виде ленты событий.

Система отслеживания ошибок (англ. *bug tracking system*) — прикладная программа, разработанная с целью помочь разработчикам программного обеспечения (программистам, тестировщикам и др.) учитывать и контролировать ошибки и неполадки, найденные в программах, пожелания пользователей, а также следить за процессом устранения этих ошибок и выполнения или невыполнения пожеланий [105].

Главный компонент такой системы — база данных, содержащая сведения об обнаруженных дефектах. Эти сведения могут включать в себя:

- номер (идентификатор) дефекта;
- короткое описание дефекта;
- кто сообщил о дефекте;
- дата и время, когда был обнаружен дефект;
- версия продукта, в которой обнаружен дефект;
- серьёзность (критичность) дефекта и приоритет решения;

- описание шагов для выявления дефекта (воспроизведения неправильного поведения программы);
- ожидаемый результат и фактический результат;
- кто ответственен за устранение дефекта;
- обсуждение возможных решений и их последствий;
- текущее состояние (статус) дефекта;
- версия продукта, в которой дефект исправлен.

Существуют различные сервисы, которые имеют множество возможностей в работе над проектами. Некоторые из них представлены в таблице 1.1.

Таблица 1.1 - Возможности сервисов

№ п/п		Мега план	План Фикс	Битр икс24	A2Б	Pyrus	JIRA	TimeM aster	vDe sk24	Risk Gap	Action space	Интра ворк	OnlyO ffice	Team Bridge	Nea ktor	MS Project
1	Управление проектом	да	да	да	да	да	да	да	да	да	да	да	да	да	да	да
2	Инструменты совместной работы	да	да	да	да	да					да	да		да	да	
3	Планирование		да		да			да						да	да	
4	Техническая поддержка и управление заявками		да									да				
5	Электронный документооборот				да	да										
6	Управление идеями				да											
7	Повышение продуктивности					да					да					
8	Отслеживание задач						да			да						
9	Управление Agile-проектом						да									
10	Учёт времени						да	да								
11	Управление персоналом								да							
12	Базы знаний									да						
13	Пакеты приложений												да			

Программа для управления проектами *Microsoft Project Server* обеспечивает следующие функции:

- управление задачами;
- отчеты и бизнес-аналитика;
- общий доступ к документам;
- управление ресурсами;
- синхронизация задач посредством SharePoint;
- управление и рассылка расписаний;
- управление запросами, проблемами и рисками;
- выбор и оптимизация портфеля;
- визуальный оптимизатор ресурсов;
- управление временем [28].

JIRA - веб-система для решения вопросов и отслеживания ошибок программного обеспечения с расширенными возможностями и функциями по управлению проектами.

JIRA помогает команде обмениваться информацией и легко вовлекать разных сотрудников в проекты и задачи, отслеживать и фиксировать ошибки пользователей в работе с программными продуктами, обеспечивать соблюдение работы точно в срок и в рамках регламента рабочего процесса, проверять и планировать эффективность работников и назначать им определённые задачи, а также отслеживать прогресс и обновление каждой задачи проектной команды.

Основные возможности JIRA:

- отслеживание и работа с проектами;
- планирование рабочего процесса;
- отслеживание эффективности [22].

В настоящее время для активного хранения проектов по разработке программно-аппаратных комплексов созданы инструментальные системы, которые называются системами управления версиями. На данный момент существуют разные типы систем, каждая из которых значима при конкретных условиях их применения.

1.3 Обзор средств организации проектной информации. Понятие системы контроля версий

Система контроля версий (Version Control System или Revision Control System) является программным обеспечением, с помощью которого обеспечивается работа с постоянно изменяемыми документами, с информацией о разработчиках и времени изменений, о прежних вариантах документов.

Данные системы требуются в работе над большими проектами, связанными с созданием программного обеспечения, для того чтобы сохранять первичный код разрабатываемых программ. Посредством системы контроля версий участники проекта могут пользоваться любыми вариантами рабочего продукта, при необходимости «вытаскивать» его старые релизы из данной системы. Поскольку система сохраняет помимо всей структуры проекта также и данные вариантов каждого документа проекта, то ее чаще называют проектным репозиторием. Часто СКВ используются в автоматизированном проектировании программно-аппаратных комплексов (САПР).

При выполнении проектных работ рабочими группами не совсем удобно выполнять мелкие изменения в проекте посредством правки главной версии продукта. В связи с этим внутри репозитория создаются ветви (**branch**) (то есть «отхождение» от главной версии) в каком-то варианте нового проекта, работа в котором проходит параллельно с изменениями основной версии. Это полезно для хранения стабильной или выпущенной версии ветви, при том, что работа над рабочим вариантом продукта продолжается. Проект может иметь несколько ветвей, параллельных работ, каждая из которых может иметь различные метки (часто называемых тэгами) для упрощения поиска. Важно при обозначении частных релизов или новых рабочих версий, созданных для внедрения. В каждой ветви существует определенное количество точек из основной версии и точек «наперед», то есть возможных вариантов кода, которые можно проверить либо «откатить» назад, либо сделать основной веткой.

Безусловно, использование системы контроля версий является обязательным

инструментом при проектировании программного обеспечения. В большинстве случаев системы контроля версий применяют централизованную модель, когда есть единый «склад» документов, который управляется выделенным сервером, выполняющим основную часть работы по управлению версиями. Перед работой пользователь копирует нужную ему версию продукта к себе на компьютер, тем самым создается так называемая локальная копия документа или «рабочая копия». Это может быть последняя версия продукта либо другая из предыдущих, которую разработчик выбирает по номеру версии или по дате создания или по иным признакам. После внесения нужных изменений в документ, новая версия продукта сохраняется «на складе». В данном случае предыдущий вариант продукта не удаляется, как в простом сохранении файлов, а также остается «на складе» и может быть взят оттуда в любое время. Сервер может использовать т. н. дельта-компрессию - такой специфичный способ хранения документов, при котором сохраняются только изменения между последовательными версиями, что позволяет уменьшить объём хранимых данных [106].

В одном проекте участвуют одновременно несколько разработчиков. Без распараллеливания работ при изменении одного и того же файла не все исправления могут быть сохранены. Системы контроля версий имеют определенный список решений такого типа ситуаций. Очень часто эти системы автоматически объединяют такого рода изменения, выполняемые разными пользователями одной команды.

Большинство систем управления версиями выделяют, помимо вышеуказанных, иные варианты использования:

- создают множество версий текущего рабочего файла, именуемыми ветками, которые сохраняют единую информацию до момента ветвления, а после нее – различную;
- хранят данные по всем изменениям документов, пользователям кто их внес и датам исправлений;
- корректируется список исправлений документов, где разработчики размещают записи и комментарии по каждому варианту файла [106].

1.4 Непрерывная интеграция

Помимо средств хранения проекта возникает необходимость слияния действий разработчиков, автоматизации большинства процессов разработки программно-аппаратных комплексов: определения данных и манипулирования ими в базе данных, осуществление проверки, просмотра кода, развертывания программного обеспечения. Члены проектной группы должны контактировать между собой для анализа состояния работы над программным обеспечением.

Возможность объединения действий разработчиков позволяет уменьшить риски, которые могут возникнуть при построении приложений, к примеру, от не высокого качества кода и обнаруженных через долгое время недостатков.

Автоматизация всех этих процессов получила название непрерывная интеграция (CI, англ. *Continuous Integration*). В ходе разработки программного кода после каждого его изменения указанные процессы включаются в систему CI.

Проект, в котором задействован один человек с несколькими сторонними связями, содержащий интеграцию программного обеспечения – не самая важная проблема. Когда проект сложный (уже при добавлении одного человека) появляется необходимость в интеграции и контроле согласованной работы составляющих программного обеспечения. Непрерывная интеграция снижает подобные риски.

В своей известной статье “*Continuous Integration*” Martin Fowler [8] описывает CI так:

«Практика разработки программного обеспечения, когда участники группы осуществляют частую интеграцию своих работ. Обычно каждый человек проводит интеграцию по крайней мере ежедневно, что приводит к нескольким интеграциям в день. Для максимально быстрого обнаружения ошибок каждая интеграция осуществляется автоматизированно (вместе с проверкой). Многие группы находят, что данный подход позволяет значительно уменьшить проблемы интеграции и способствует более быстрой разработке программного обеспечения.

А это означает следующее:

- все разработчики выполняют закрытое построение на собственных рабочих станциях перед передачей кода в хранилище с контролем версий, для гарантии того, что внесенные изменения не приведут к ошибке при интеграционном построении;
- разработчики обновляют свой код в хранилище с контролем версий *по крайней мере* один раз в день;
- интеграционное построение осуществляется несколько раз в день на выделенной для этого машине;
- при каждом построении проводится 100 % проверок;
- создаваемый продукт пригоден для функциональной проверки;
- исправление ошибок имеет самый высокий приоритет;
- часть разработчиков просматривают отчеты, созданные в ходе построения, стандарты программирования и отчеты анализа зависимостей, в поисках областей для усовершенствования ».

За счет непрерывной интеграции увеличивается вероятность обратной связи, которая дает возможность в течение дня отслеживать работу над проектом. *CI* используется для снижения временного промежутка между временем появления ошибки и ее исправлением. В связи с этим программное обеспечение в целом становится наиболее качественное.

Постоянное использование *CI* предоставляет информацию по общему состоянию исполняемого программного обеспечения. Кроме этого, непрерывная интеграция работает с рефакторингом и методом проверки, так как они основаны на идеи небольших изменений. По факту, *CI* обеспечивает сочетание последних с оставшимся пакетом программного обеспечения. За счет снижения объема ручного труда в процессе редактирования исполняемого программного обеспечения *CI* увеличивает совместную ответственность рабочей команды и уменьшает трудозатратность проекта [95].

1.5 Описание автоматизированной системы отдела технической документации на примере крупной проектной организации ФНПЦ АО «НПО Марс»

Объекты автоматизации [98]

- Картотеки архивов. Содержат реквизиты документов, хранящихся в архиве. Для архивов электронных документов одним из реквизитов является ссылка на файлы документа, находящиеся в файловом хранилище. Объединяет девять картотек основных архивов с различными наборами реквизитов и правилами учета и хранения: электронные подлинники и дубликаты, бумажные подлинники и дубликаты, покупное программное обеспечение, сопроводительная документация на изделия предприятия и покупные изделия, документация научно-технического задела, промежуточная проектная и эскизная документация, бумажные копии для обеспечения производства. Обеспечивает учет документации в соответствии с ГОСТ 2.501, ГОСТ 2.503, ГОСТ 2.102, ГОСТ 28388, ГОСТ 2.601. Обеспечивают учет документов по предварительным извещениям.
- Архивное файловое хранилище. Содержит все версии электронных документов, когда-либо зарегистрированных в архивах; скан-копии ряда бумажных документов; файлы промежуточных редакций документов, направленных на предварительную проверку. Имеет иерархическую структуру, использует накопительный принцип заполнения, большая часть хранилища заполняется автоматически при работе с картотеками системы. Целостность хранилища обеспечивается автоматическим созданием контрольных копий документации.
- Файловое хранилище общего пользования. Содержит исходные файлы актуальных версий электронных документов, а также скан-копии бумажных документов. Также имеет иерархическую структуру, заполняется полностью автоматически при регистрации документов в

картотеках. Используется для обеспечения электронными версиями документов всех сотрудников предприятия в соответствии с системой разграничения доступа.

- Система внутреннего документооборота отдела технической документации. Содержит картотеки внутренних объектов, необходимых для работы отдела, обеспечивает движение документов по участкам работ, контроль и учет выполнения работ.
- Система разграничения доступа к документам.
- Объекты, обеспечивающие интеграцию с автоматизированными системами других подразделений предприятия.

Функции системы

Функции архивной подсистемы

- Автоматизированная, либо автоматическая регистрация документов в картотеках архивов, автоматическое формирование инвентарных книг по ГОСТ 2.501, исключение возможности дублирования данных, в том числе в режиме многопользовательской работы. Автоматическая классификация документов при регистрации.
- Автоматический расчет контрольной суммы для обеспечения подтверждения неизменности хранимых файлов электронных документов.
- Автоматическое заполнение файловых хранилищ при регистрации электронных документов.
- Автоматизированная предварительная проверка документов с загрузкой данных из файлов реквизитов.
- Автоматический подбор комплекта документов для создания копий.
- Автоматическое создание сопроводителей к комплекту документов по ГОСТ 28388.
- Автоматическая регистрация скан-копий бумажных документов и размещение их в файловом хранилище.

- Автоматизированная регистрация и учет применяемости документов с формированием структуры документации на изделие.
- Автоматизированная проверка комплектности документации на изделие по головной спецификации.
- Реквизитный поиск по картотекам отдельных архивов, либо общий поиск по всей системе.
- Контроль актуальности версий используемых (изготавливаемых) копий документов.

Функции подсистемы электронного документооборота ОТД

- Ведение внутренних справочников системы с возможностью отслеживания «похожих» данных.
- Автоматизированная регистрация выполненных работ по учету и изготовлению документации на участках работ. Индивидуальный учет выполненных работ, формирование карты учета рабочего времени.
- Регистрация входящих заявок на выполнение работ и распределение работ по участкам. Временной контроль выполнения работ.

Функции подсистемы разграничения доступа

- Ведение классификационных справочников документации, справочников пользователей и структурных подразделений предприятия.
- Ведение карты предоставления доступа сотрудников к документации электронного архива. Предоставление доступа на основании классификационных данных документов и структурной организации предприятия. Предоставление доступа на основе принципа «дружественных» подразделений.
- Обеспечение возможности получения информации о корректировках документов – подписка на обновления документа.
- Обеспечение возможности оформления доступа к документам, еще не зарегистрированным в архиве – на основе классификационных данных.

- Сбор и обработка статистики обращений сотрудников предприятия к данным архивов.

Принципиальные отличия от существующих систем архивов и электронного документооборота

- Ведение картотек архивов с полным учетом требований ГОСТ ЕСКД и ЕСПД.
- Обеспечение возможности формирования сопроводительных и отчетных документов на бумажных носителях в соответствии с ГОСТ.
- Индивидуальный учет выполненных работ с возможностью контроля и формирования общего отчета.
- Возможность получения данных для анализа «узких» мест при выполнении работ по обработке и изготовлению документации.
- Разграничение прав доступа к документу на основании классификации документации и структурной организации предприятия. Возможность оформления доступа к еще не зарегистрированным документам.



Рисунок 1.4 – Электронные картотеки

В картотеке ведется учет изменений каждого документа - к каждому изменению прикреплены файлы текущей версии, изменение регистрируется только окончательно утвержденное - выпущенное по извещению.

Система, обеспечивающая автоматизацию этапов разработки проекта и управление данными о самом изделии в частности, которые используются на конструкторском предприятии (PDM-система), включает в себя систему контроля версиями и занимает одно из главных мест при реализации проекта организации.

Одним из требований при реализации любого типа проекта является наличие большого количества отчетности в виде документации, требующего анализа и правильного управления для увеличения эффективности и во избежание рисков.

Проектный документ, хранящийся в архиве, представляет собой набор файлов, содержание которых выполнено с учетом требований определенных

стандартов. При разработке документа ему заранее присваивается десятичный номер. Каждый разрабатываемый документ в процессе согласования и сдачи в архив сопровождается листом утверждения (ЛУ) и информационно-поисковой характеристикой (ИПХ). В ЛУ проставляются подлинные подписи лиц, разработавших, проверивших и согласовавших документ. ЛУ является самостоятельным документом и хранится на учете в архиве бумажных подлинников. Также ИПХ содержит все реквизиты документа, является первичным документом при заполнении картотеки и хранится в архиве в бумажном виде как неотъемлемая часть документа.

Для организации быстрого доступа к реквизитам документов существует картотека, каждый документ в которой представлен отдельной карточкой с записанными в ней реквизитами (аналог библиотечного каталога). Карточки упорядочены по одному какому-либо признаку, выбранному архивариусом и однозначно определяющему документ [133]. На данном предприятии таким признаком является десятичный номер (обозначение) документа. В карточку также записываются следующие обязательные реквизиты документа:

- инвентарный номер;
- номера извещений;
- автор (подразделение-разработчик);
- даты корректировок документа;
- контрольные характеристики файлов, входящих в состав документа;
- сведения о выданных учтенных копиях;
- первичная применяемость документа.

Для организации хранения всех версий документов, стоящих на учете в архиве, определена строгая иерархия каталогов. Верхний уровень иерархии составляют наименования изделий и тем работ, средний уровень состоит из наименований подразделений-разработчиков и третий, последний уровень составляют номера версий документов в соответствии с проведенными извещениями [85].

При постановке на учет нового документа ему присваивается инвентарный номер и заводится карточка, в которую записываются реквизиты документа. Исходя из темы (изделия), к которой принадлежит документ, и подразделения, ответственного за его разработку, архивариус определяет место документа в иерархии архива и записывает пути к месту хранения файлов в картотеку. Инвентарный номер, дата приема в архив и место хранения файлов также дублируются в ИПХ, которые хранятся в архиве в отдельных папках, упорядоченные по инвентарному номеру документа.

При постановке на учет новой версии существующего документа архивариус подбирает карточку документа и ИПХ предыдущей версии. Архивариус также определяет место хранения файлов, исходя уже не только из темы и подразделения, но и из номера версии (или изменения, что точнее с точки зрения архивариуса).

В карточку записывается номер извещения, по которому проводится корректировка документа, номер изменения, новые пути к файлам. К инвентарному номеру добавляется номер изменения документа, также в карточку добавляется дата корректировки документа. Все реквизиты дублируются в ИПХ новой версии, старая ИПХ аннулируется, на ее место в папку раскладывается ИПХ новой версии.

В отдел технической документации (ОТД) предприятия входит несколько архивов и групп учета различной документации. Связующим звеном между службами ОТД и другими подразделениями предприятия является диспетчерская ОТД, выполняющая также функцию распределения работ по различным участкам ОТД и контроля своевременности выполнения работ [90].

Система, организующая электронные архивы проектной документации позволяет предприятию в структурированном виде накапливать информацию, содержащуюся в документах. Накопленные документы можно использовать не только в повседневной работе организации. Так как наиболее часто изменяемыми документами являются разработки программного обеспечения проекта, то одной из главных возникающих задач является отслеживания появления ошибок в

результате объединения работы участников проекта. Поэтому для решения поставленных задач систему контроля версий интегрируют с сервисом управления проектами, тем самым руководители получают ряд возможностей, основными из которых являются сотрудничество со всеми участниками проекта, отслеживание эффективности, планирование рабочего процесса. Но для главного конструктора и его заместителей система управления проектами не может выдавать информацию о возможных проблемах в реализации проекта, а также выявить некоторые тенденции и составить прогнозы развития с учетом нескольких показателей деятельности работы над проектом.

Для ведения кода проекта на предприятии используется Microsoft Project Server, где формируются задачи, отслеживаются работы. Используя эту инструментальную систему, возможно построить метрики для контроля каждого проекта. Динамика метрик выражается временными рядами значений.

Прогнозирование в задачах анализа поступившей в архив документации применяется в трех случаях:

1. Прогнозирование комплектности документов (количество или процент сданных в архив в текущий период времени).
2. Прогнозирование документов по проекту.
3. Прогнозирование извещений по проекту.

Целью прогнозирования является предположения будущих значений количества документов последующих дат.

Прогнозирование поступления проектных документов подразумевает аналитико-синтетическую переработку и заключается в выборе наилучшей гипотезы прогноза поступления документов проекта.

1.6 Основные методы интеллектуального анализа временных рядов

1.6.1 Базовые понятия анализа временных рядов

Временной ряд (или ряд динамики) — последовательно измеренные через некоторые (зачастую равные) промежутки времени данные [77]. В каждом

временном ряду присутствуют два основных элемента: время и конкретное значение показателя (уровень ряда).

Тренд (от англ. trend — тенденция) — тенденция изменения показателей временного ряда. Тренды могут быть описаны различными функциями — линейными, степенными, экспоненциальными и т. д. Тип тренда устанавливают на основе данных временного ряда, путем осреднения показателей динамики ряда, на основе статистической проверки гипотезы о постоянстве параметров графика [119].

Тенденция — это некоторая закономерность, присущая анализируемому объекту, которая прослеживается с момента начала анализа, на всем протяжении и в подводимых на его основании итогах.

Прогноз — суждение о будущем состоянии объекта исследования, основанное на специальном научном исследовании (прогнозировании).

Прогнозирование — ход разработки прогноза. *Прогнозирование временных рядов* представляет собой создание модели для предсказания будущих значений на основе известных.

Модель прогнозирования — модель объекта прогнозирования, исследование которой позволяет получить информацию о возможных состояниях объекта прогнозирования в будущем и (или) путях и сроках их осуществления.

Методика прогнозирования — совокупность специальных правил и приемов (одного или нескольких методов) разработки прогнозов.

Прогнозный горизонт — максимально возможный период упреждения прогноза заданной точности.

Ошибка прогноза — апостериорная величина отклонения прогноза от действительного состояния объекта.

Предиктор (от англ. predictor «предсказатель») — прогностический параметр; средство прогнозирования. В технической области термин употребляется как независимая переменная; прогнозирующее устройство; экстраполятор; прогнозирующий параметр [96].

Анализ временного ряда объединяет методы изучения временных рядов, как пытающиеся понять природу точек данных, так и пытающиеся построить прогноз [77].

Анализ временных рядов предполагает решение конкретных задач, таких как: измерение абсолютной и относительной скорости роста, либо снижение уровня за отдельные промежутки времени; обобщение характеристик уровня и скорости его изменения за тот или иной период; выявление основных тенденций развития явлений на отдельных этапах; выявление факторов, обуславливающих изменение изучаемого явления во времени; прогнозирование развития явления в будущем (экстраполяция и интерполяция).

Определение *анализ временных рядов* применяется тогда, когда необходимо выделить текущий вид анализа от:

- 1) простейших задач анализа данных (когда нет естественного порядка поступления сведений);
- 2) анализа пространственных данных, в котором наблюдения зачастую связаны с географическим положением.

Модель временного ряда показывает насколько будут теснее связаны близкие во времени наблюдения, в отличие от удалённых. К тому же модели временных рядов очень часто применяют однонаправленный порядок по времени, то есть значения в ряду представляются через предыдущие значения, а не через будущие.

Методы анализа временных рядов классифицируются на анализ в частотной области, который строится на спектральном и вейвлетном анализе для применения в экспериментах на предварительном этапе, и анализ во временной области.

1.6.2 Методы интеллектуального анализа временных рядов

На основе новой методологии Data Mining решается расширенная совокупность задач анализа временных рядов, определенных в работах [6], [143]:

- 1) сегментация – разбиение ВР на значимые сегменты [16];
- 2) кластеризация – поиск группировок ВР или их паттернов [12];

- 3) классификация – назначение ВР или их паттернам одного из заранее определенных классов [18];
- 4) индексирование – построение индексов для эффективного выполнения запросов к базам данных ВР;
- 5) резюмирование (summarization) – формирование краткого описания ВР, содержащего существенные черты с точки зрения решаемой задачи [24];
- 6) обнаружение аномалий – поиск новых, не типичных паттернов ВР;
- 7) частотный анализ – поиск часто проявляющихся паттернов ВР [52];
- 8) прогнозирование – прогноз очередных значений ВР на основе истории ВР.

Прогнозирование – получение многоуровневого прогноза: в терминах общей, локальной, элементарной тенденций, в терминах нечетких и числовых уровней ВР.

Резюмирование – генерация гранулярного описания поведения НВР [56], определение семантики гранул и трансляция гранулярного описания НВР в виде предложения на естественном языке [55].

Поиск аномалий – выявление лексических, синтаксических и семантических ошибок и нетипичных нечетких объектов и правил [72], [125].

В прогностике существуют различные частные виды классических задач на прогнозирование. Формулирование таких задач единообразным образом позволяет сравнивать различные методы, предлагаемые различными дисциплинами. Развитие методов прогнозирования непосредственно связано с развитием информационных технологий, в частности, с ростом объемов хранимых данных и усложнением методов и алгоритмов прогнозирования, реализованных в инструментах Data Mining [122].

Задача прогнозирования, может считаться одной из наиболее сложных задач Data Mining, она требует тщательного исследования исходного набора данных и методов, подходящих для анализа [71].

Прогнозирование (forecasting) является одной из задач Data Mining и одновременно одним из ключевых моментов при принятии решений.

Data Mining – мультидисциплинарная область, возникшая и развивающаяся на базе таких наук как прикладная статистика, распознавание образов, искусственный интеллект, теория баз данных и др., см. рис. 1.5.

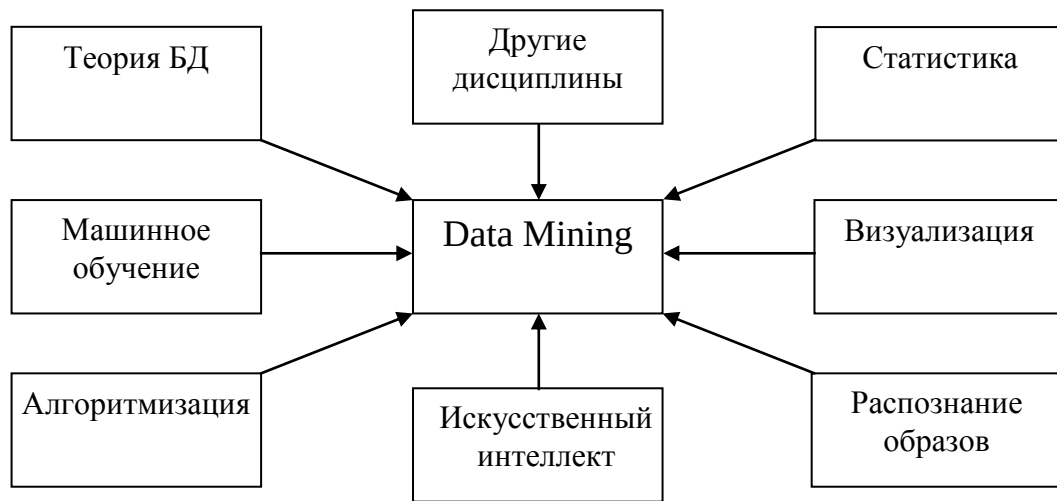


Рисунок 1.5 – Data Mining как мультидисциплинарная область

Решение задачи прогнозирования требует некоторой обучающей выборки данных.

В самых общих чертах решение задачи прогнозирования сводится к решению таких подзадач:

- выбор модели прогнозирования;
- анализ адекватности и точности построенного прогноза.

Выбор метода прогнозирования зависит от многих факторов, в том числе от параметров прогнозирования. Выбор метода следует производить с учетом всех специфических особенностей набора ретроспективных данных и целей, с которыми он строится.

Программное обеспечение Data Mining, используемое для прогнозирования, должно обеспечивать пользователя точным и достоверным прогнозом. Однако получение такого прогноза зависит не только от программного обеспечения и методов, заложенных в его основу, но также и от других факторов, среди которых полнота и достоверность исходных данных, своевременность и оперативность их пополнения, квалификация пользователя [122].

Самыми простыми методами прогнозирования являются "наивные" модели прогнозирования, При создании которых предполагается, что некоторый

последний период прогнозируемого временного ряда лучше всего описывает будущее этого прогнозируемого ряда, поэтому в этих моделях прогноз, как правило, является очень простой функцией от значений прогнозируемой переменной в недалеком прошлом.

Самой простой моделью, основанной на простом усреднении является «Средние» и «скользящие средние» и в отличие от самой простой "наивной" модели, которой соответствовал принцип "завтра будет как сегодня", этой модели соответствует принцип "завтра будет, как было в среднем за последнее время". Такая модель, конечно более устойчива к флуктуациям, поскольку в ней сглаживаются случайные выбросы относительно среднего. Несмотря на это, этот метод идеологически настолько же примитивен, как и "наивные" модели и ему свойственны почти те же самые недостатки.

Одним из наиболее популярных методов решения задач прогнозирования является метод деревьев решений (decision trees). На сегодняшний день существует большое число алгоритмов, реализующих деревья решений.

Так же задача прогнозирования решается с помощью нейронных сетей. Для нейронной сети задача прогнозирования может быть поставлена таким образом: найти наилучшее приближение функции, заданной конечным набором входных значений (обучающих примеров). Например, нейронные сети позволяют решать задачу восстановления пропущенных значений [145].

1.7 Нечеткие модели временного ряда

1.7.1 Основы нечеткого моделирования временных рядов

По сравнению с устоявшимся ВР значениями нечеткого ВР (НВР) являются нечеткие множества, а не фактические величины состояния процесса. В 1993 году Сонг и Чиссом [47] предложили модели стационарных и нестационарных (time-invariant и time-variant) НВР первого порядка (*first-order*) и применили разработанные модели для прогнозирования количества регистрирующихся студентов университета штата Алабама, фазифицировав предварительно четкий

ВР. Это было первое применение нечетких моделей в моделировании ВР и первое определение моделей НВР [144].

Допустим, дан процесс, состояния которого описываются n значениями одной переменной. В результате наблюдения получен ВР этой переменной, представляющий последовательность упорядоченных в равноотстоящие моменты времени пар $\{x_i, t_i\}$, таких, что $\forall x_i \in X, X \subset R^1, t_i \in N, i \in [1, n]$. Значение x_i называют *уровнем* ВР [67].

Нечетким временным рядом (НВР) называют упорядоченную последовательность наблюдений над некоторым событием, состояния которого изменяются во времени, если значение состояния в момент t_i выражено с помощью нечеткой метки $\tilde{x}_i \in \tilde{X}, i \in [1, n]$, n – количество членов ряда. Таким образом, НВР можно представить в виде $\tilde{Y} = \{\tilde{x}_i, t_i\}$, где $\tilde{x}_i$ – i -тое нечеткое множество (нечеткая метка), t_i – i -тое значение момента времени, $t_1 \leq t_i \leq t_n$, n – количество членов НВР [145].

Нечеткая метка – это понятие на естественном языке, получаемое посредством фазификации исходного четкого ВР [69]. Следовательно, для каждого четкого элемента ВР существует пара $(\mu_{\tilde{x}_i}, \tilde{x}_i)$, где $\tilde{x}_i$ – лингвистическое понятие (нечеткая метка); $\mu_{\tilde{x}_i} \in [0, 1]$ – функция принадлежности нечеткой метки $\tilde{x}_i$ уровню ВР x_i .

Целью нечеткого моделирования числового ВР $Y = x_t, x_t \in X, t = 1, 2, \dots, n$ является идентификация нелинейной модели [93], описывающей динамику исследуемого процесса, в виде модели НВР p -го порядка [46], [42]:

$$\tilde{X}_t = (\tilde{X}_{t-1} \times \tilde{X}_{t-2} \times \dots \times \tilde{X}_{t-p})^\circ R(t, t-p),$$

$$R(t, t-p) = \cup_{i,j} R_{i,j}(t, t-p) = \max_p \left\{ \min_{j, i_1, i_2, \dots, i_p} \{ \tilde{x}_t^j, \tilde{x}_{t-1}^{i_1}, \dots, \tilde{x}_{t-p}^{i_p} \} \right\},$$

где $\tilde{X}_i$ – нечеткие множества (функции принадлежности), заданные на X ; $R(t, t-p)$ – система нечетких отношений; $^\circ$ – операция *max-min* композиции.

Данная модель имеет лингвистическую интерпретацию в виде базы нечетких правил «Если-То»:

$R(t, t - p)$: Если x_{t-p} есть $\tilde{x}_{t-p}$, x_{t-p+1} есть $\tilde{x}_{t-p+1}, \dots, x_{t-1}$ есть $\tilde{x}_{t-1}$,
то x_t есть $\tilde{x}_t$.

При моделировании числовых ВР в рамках нечеткого динамического процесса числовые значения предварительно преобразуют в нечеткие (фазифицируют)

$$\tilde{x}_t = Fuzzy(x_t),$$

далее результаты дефазифицируют

$$x_t = deFuzzy(\tilde{x}_t).$$

1.7.2 Обзор методов нечеткого моделирования временных рядов

Главной задачей при построении модели динамического процесса в виде НВР является идентификация его модели по наблюдениям $X = (x_t), t = 1, 2, \dots, n$ в форме нечетких правил «Если-То» [37], [135]. Эти правила используются для «вычисления» приближенных значений $\hat{X} = (\hat{x}_t), t = 1, 2, \dots, n$.

Условия применимости алгоритма нечеткого моделирования числового ВР:

1. Задать значения ВР: $X = (x_t), t = 1, 2, \dots, n$.
2. Определить универсальное множество $U = \{nmin, nmax\}, U \supset X$, параметры и набор функций принадлежности нечётких множеств, определенных на этом универсальном множестве $\tilde{X} = (\tilde{x}_i), i = 1, 2, \dots, m$.

Моделирование НВР в соответствии с нечеткой моделью, предложенной в работе [47], состоит в реализации следующих шагов.

1. Фазификации значений ВР $\tilde{x}_t = Fuzzy(x_t), t = 1, 2, \dots, n$.
2. Формирование отношения вида $R_{t,t-1}: \tilde{x}_{t-1} \rightarrow \tilde{x}_t, t=2, 3, \dots, n$. Так как это выражение соответствует нечеткой импликации, то его можно представить в лингвистической форме $R(t, t - 1)$: Если x_{t-1} есть $\tilde{x}_{t-1}$, то x_t есть $\tilde{x}_t$.

Модель НВР, по Сонгу, может быть представлена в форме уравнения:

$$\tilde{X}_t = \tilde{X}_{t-1} \circ R(t, t - 1).$$

Реализация нечеткого логического вывода по данной модели выполняется по алгоритму Мамдами [86], [145]:

1. Присвоить $t=2$.

2. Пока $t \leq n$ исполнять:

а) фаззифицировать входные данные – определение степени принадлежности входных данных входным нечетким переменным. Подставить значения x_{t-1} в левые части нечетких правил.

б) Вычислить результат применения нечеткого правила $R_{ij}(t, t-1)$ для каждой импликации $R_{t,t-1}: \tilde{x}_{t-1} \rightarrow \tilde{x}_t, t = 2, 3, \dots, n$.

в) Провести агрегацию правых частей нечетких правил $\tilde{x}_t = \max(\tilde{x}_j), j = 1, 2, \dots, r$.

г) Вычислить приближенное значение числового ВР

$$\hat{x}_t = \frac{\int_{n \min}^{n \max} x \cdot \tilde{x}_t(x) dx}{\int_{n \min}^{n \max} \tilde{x}_t(x) dx}.$$

е) Перейти к оцениванию следующего члена ВР $t = t+1$ [20].

Предложенная Сонгом [47] модель НВР имеет следующие недостатки:

1. Экспертное задание количества входных, выходных переменных и параметров нечетких множеств.
2. При реализации нечеткой композиции (max-min) модели, требуется большое количество вычислений, особенно, когда нечеткое отношение очень велико.
3. Отсутствуют возможности проверки на полноту и непротиворечивость базы правил и поиска наилучшей модели, что приводит к недостаточности точности модели.
4. Не реализован анализ нечетких тенденций ВР.
5. Отсутствуют критерии качества нечеткого моделирования.

1.7.3 Нечеткая тенденция. Виды НТ (элементарная, локальная, общая)

Нечеткой тенденцией (НТ) нечеткого временного ряда будем называть нечеткую метку, выражающую характер изменения (систематическое движение) последовательности нечетких уровней НВР в заданном интервале времени. Нечеткая тенденция выражает поведение НВР в лингвистическом виде, например: «Рост», «Падение», «Стабилизация», «Колебания», «Хаос». Для нечетких термов, обозначающих тенденцию, возможно применение модификаторов «очень», «более-менее» и т. д. Отметим важное свойство нечетких экспертных оценок, обусловленное возможностью их ранжирования, что позволяет представить их совокупность в виде некоторой системы (шкалы) с отношениями. Бинарные отношения, образованные на множестве нечетких экспертных оценок, порождают сравнительные оценки по различным критериям, такие, как «Больше», «Меньше», «Примерно Равны», «Рост», «Падение», «Предпочтительнее», «Лучше». Такие сравнительные оценки представляют изменения (различие) нечетких меток в различных пространствах: в пространстве объектов, во временном пространстве, в пространстве задач и характеризуют тенденции.

Изменения нечетких меток во временном пространстве порождают нечеткий временной ряд с нечеткой тенденцией.

Содержательно термин «тенденция» ассоциируется с изменением и используется в предложениях совместно с общеупотребительными лингвистическими оценками изменений, связанными с протяженностью, типом и интенсивностью, например, «длительная тенденция роста», «ярко выраженная тенденция падения», «тенденция стабильного высокого качества» и т. д [66], [67].

Выделяют [126] следующие свойства нечеткой тенденции (НТ):

- А. Нечеткость. Это свойство обозначает факт, что НТ построена на основе нечетких значений НВР и наследует нечеткость этих значений.
- В. Протяженность. Это свойство связано с различной длительностью НТ.

- С. Типичность. Свойство типичности НТ позволяет различать классы, типы НТ, внутри которых нечеткие тенденции будут рассматриваться как однородные.
- Д. Значимость. Для различия нечетких тенденций одного типа и одинаковой длительности целесообразно использовать характеристику степени значимости или интенсивности НТ.
- Е. Ориентированность во времени. Это свойство обозначает, что нечеткие тенденции определяются в направлении увеличения временных отсчетов.
- Ф. Лингвистическая интерпретируемость. Данное свойство нечеткой тенденции следует из определения нечеткой тенденции. По определению, НТ есть нечеткая метка, которой сопоставляется лингвистически терм.

Формальным аппаратом для извлечения знаний из НВР, который позволяет изображать поведение ВР в терминах лингвистических оценок тенденций, являются ACL-шкала (absolute comparative linguistic) и ее функциональное наполнение, связанное с генерацией сравнительных нечетких оценок изменений [68].

Рассмотрим виды ACL-шкал и нечетких оценок изменений, получаемых на этих шкалах с целью формализации свойств нечеткой тенденции нечеткого временного ряда.

Так, в случае номинальной ACL-шкалы нечеткая тенденция задается бинарной оценкой $E_0 = (\text{Есть изменения, Нет изменений})$.

Для ранговой ACL-шкалы определены понятия БОЛЬШЕ, МЕНЬШЕ, РАВНО, которые задают тип тенденции. Поэтому, если ACL-шкала ранговая, то тенденция может быть выражена в нечетких оценках, отражающих тип тенденции $E_1 = (\text{Нет Изменений, Есть Изменения (в сторону увеличения, в сторону уменьшения)})$ или $E_1 = (\text{Стабильность, Рост, Падение})$.

В том случае, если имеется возможность определить степень выраженности типа изменения, логика оценок уточняется за счет включения этого

дополнительного параметра, для оценки которого должна быть определена своя шкала, например, бальная, ранговая или числовая. Тогда результат «измерения» на ACL-шкале изменений может быть представлен в виде $E2 = (\text{Стабильность}, \text{Рост}(\tilde{A}), \text{Падение}(\tilde{A}))$, где $\tilde{A}$ – оценки для «измерения» степени интенсивности тенденции.

В более общем случае каждая оценка $E2$ представима следующим образом: $((v_k(p_{ks}))$, где v_k – тип изменения, k – количество типов изменения, p_{ks} – параметры типов изменения, s – количество параметров. Следует отметить, что при оценивании двух нечетких уровней нечеткого временного ряда каждая оценка $E0, E1, E2$ наследует свойства нечеткости этих значений, то есть вычисленную при фаззификации степень принадлежности [67].

Авторы [126] отмечают особенности оценивания тенденций в базисе оценок $E2$ для последовательности нечетких значений, представленных в виде нечеткого временного ряда (НВР).

Первая особенность заключается в том, что исходный НВР порождает временной ряд степеней принадлежности нечетких тенденций.

Вторая особенность: при использовании ACL-шкал для оценивания тенденций в последовательности нечетких значений НВР [76], результаты формируются в виде множеств, соответствующих типам v_k оценок $E2$. Каждое из этих множеств характеризуется функцией распределения выраженности α_{v_k} и мощностью, которая в простейшем случае сопоставима с длительностью k -того типа нечеткой тенденции Δt_{v_k} . Указанные функции $\alpha_{v_k}, \Delta t_{v_k}$ образуют параметры типов тенденций, а последовательность типов тенденций определяет структуру тенденции НВР.

Следовательно, при определении тенденции НВР на основе оценок $E0, E1, E2$ модель нечеткой тенденции τ_k может быть представлена с разной степенью детализации:

$$\tau_k(\tilde{X}, E0) = (v_k, \mu_k),$$

$$\tau_k(\tilde{X}, E1) = (v_k, \mu_k),$$

$$\tau_k(\tilde{X}, E2) = (v_k, \mu_k, \alpha_{v_k}, \Delta t_{v_k}),$$

где τ_k – наименование нечеткой тенденции;

v_k – тип нечеткой тенденции, например, для оценки E0 из базового множества типов {«Есть изменения», «Нет изменения»}, для оценок E1 и E2 из базового множества {«Рост», «Падение», «Стабильность»}. Базовое множество может быть расширено за счет производных типов тенденций;

μ_k – функция принадлежности нечеткой тенденции, которая формируется на основе образующих ее нечетких значений ВР;

α_{v_k} – интенсивность каждого типа нечеткой тенденции, контекстное расширение типа тенденции;

Δt_{v_k} – длительность каждого типа нечеткой тенденции.

Модель нечеткой тенденции НВР рассматривают [126] в виде оценок E2 как наиболее информативную.

Нечеткие метки «Рост», «Падение», «Стабильность» называют базовыми типами тенденций, а «Колебания», «Хаос» – называют производными типами, так как вывод относительно их типа формируется на основе базовых.

Тогда нечеткое множество ACL-шкалы $\tilde{V}$ может быть использовано для обозначения типов тенденций НВР. Представим $\tilde{V} = \tilde{V}1 \cup \tilde{V}2$, где $\tilde{V}1$ определяет множество базовых типов тенденций НВР, которое конечно и ограничено тремя типами $\tilde{V}1 = \{\tilde{v}1, \tilde{v}2, \tilde{v}3\}$, например, $\tilde{v}1$ = «Стабилизация», $\tilde{v}2$ = «Падение», $\tilde{v}3$ = «Рост», $\tilde{V}2$ – множество производных типов тенденций НВР, которое также конечно и ограничено, например, $\tilde{V}2 = \{\tilde{v}4, \tilde{v}5\}$, где $\tilde{v}4$ = «Колебания периодические», $\tilde{v}5$ = «Колебания хаотические».

Пусть нечеткое множество $\tilde{A}$ для заданного нечеткого временного ряда генерируется в процессе построения ACL-шкалы автоматизированной процедурой.

Для построения лингвистических оценок протяженности тенденции строят универсальное множество ΔT , такое, что $\forall \Delta t \in \Delta T$. На этом универсуме

определяют ACL-шкалу $S_{\Delta T}$ для генерации лингвистической нечеткой оценки длительности нечеткой тенденции $\Delta\tilde{t} \in \Delta\tilde{T}$.

Согласно определению нечеткой тенденции [145], каждая нечеткая тенденция τ_k нечеткого временного ряда $\tilde{Y} = \{t_i, \tilde{x}_i\}, i \in [1, n], \tilde{x}_i \in \tilde{X}$ может быть представлена нечеткой меткой, выражающей характер поведения НВР.

В рамках теории нечетких множеств нечеткая тенденция представима нечетким множеством $\tau_k = \{\tilde{X}, \mu_{\tau_k}(\tilde{X})\}, k \in [1, n]$, где $\mu_{\tau_k}(\tilde{X})$ – функция принадлежности нечеткой тенденции τ_k , определенная на множестве нечетких уровней $\tilde{X}$ нечеткого временного ряда.

Каждая нечеткая тенденция $\tau \in \mathfrak{S}$ нечеткого временного ряда $\tilde{Y} = \{t_i, \tilde{x}_i\}$ может быть представлена структурной моделью в виде кортежа отношения, построенного на декартовом произведении свойств нечеткой тенденции $\tilde{V} \times \tilde{A} \times \Delta\tilde{T} \rightarrow \mathfrak{S}$:

$$\tau = \langle \tilde{v}, \tilde{a}, \Delta\tilde{t}, \mu \rangle,$$

где τ – наименование нечеткой тенденции (идентификатор) из множества $\mathfrak{S}$, $\tau \in \mathfrak{S}$;

$\tilde{v}$ – тип нечеткой тенденции (тип изменений), $\tilde{v} \in \tilde{V}$;

$\tilde{a}$ – степень изменения, то есть интенсивность нечеткой тенденции, $\tilde{a} \in \tilde{A}$;

$\Delta\tilde{t}$ – продолжительность нечеткой тенденции, $\Delta\tilde{t} \in \Delta\tilde{T}$;

μ – функция принадлежности участка нечеткого временного ограниченного $\Delta\tilde{t}$ нечеткой тенденции τ .

При значении функции принадлежности, равном единице, нечеткая тенденция относится к четким тенденциям.

Однородными нечеткими тенденциями считаются нечеткие тенденции $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta\tilde{t}_i, \mu_i \rangle$ и $\tau_s = \langle \tilde{v}_s, \tilde{a}_s, \Delta\tilde{t}_s, \mu_s \rangle$, для которых верно $\tilde{v}_i = \tilde{v}_s$.

Объединением однородных нечетких тенденций $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta\tilde{t}_i, \mu_i \rangle$ и $\tau_s = \langle \tilde{v}_s, \tilde{a}_s, \Delta\tilde{t}_s, \mu_s \rangle$ является $\tau_j = \tau_i \cup \tau_s$, такая что $\tau_j = \langle \tilde{v}_j, \tilde{a}_j, \Delta\tilde{t}_j, \mu_j \rangle$, для которой $\tilde{v}_j = \tilde{v}_s$, $\tilde{a}_j = \text{Union}(\tilde{a}_i, \tilde{a}_s)$, $\mu_j = \mu_i \cup \mu_s$, длительность тенденции $\Delta\tilde{t}_j = \text{Union}(\Delta\tilde{t}_i, \Delta\tilde{t}_s)$.

Эквивалентными нечеткими тенденциями являются однородные нечеткие тенденции $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta\tilde{t}_i, \mu_i \rangle$ и $\tau_s = \langle \tilde{v}_s, \tilde{a}_s, \Delta\tilde{t}_s, \mu_s \rangle$, такие, что $\tilde{a}_i = \tilde{a}_s$, $\Delta\tilde{t}_i = \Delta\tilde{t}_s$, $\mu_i = \mu_s$.

Неоднородными нечеткими тенденциями считают нечеткие тенденции $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta\tilde{t}_i, \mu_i \rangle$ и $\tau_s = \langle \tilde{v}_s, \tilde{a}_s, \Delta\tilde{t}_s, \mu_s \rangle$, такие, что $\tilde{v}_i \neq \tilde{v}_s$.

Подобными нечеткими тенденциями являются однородные нечеткие тенденции $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta\tilde{t}_i, \mu_i \rangle$ и $\tau_s = \langle \tilde{v}_s, \tilde{a}_s, \Delta\tilde{t}_s, \mu_s \rangle$ такие, что $\tilde{a}_i = \tilde{a}_s$, $\Delta\tilde{t}_i = \Delta\tilde{t}_s$.

Семантика введенных компонент нечеткой тенденции и операций их объединения, сравнения определяется их нечетко-множественной природой и моделируется на основе отдельных ACL-шкал.

Рассмотрим нечеткий временной ряд $\tilde{Y} = \{t_i, \tilde{x}_i\}, i \in [1, n]$, n – количество членов нечеткого ВР. В том случае, если модель нечеткой тенденции определяется для двух соседних уровней НВР, такую нечеткую тенденцию τ отнесем к классу *элементарных нечетких тенденций* $\tau \in \mathfrak{Z}$, если модель нечеткой тенденции определяется между начальным и конечным уровнем НВР, такую нечеткую тенденцию авторы [126] классифицируют как *общую (основную) тенденцию* $G\tau \in G\mathfrak{Z}$ НВР, в противном случае, нечеткие тенденции, определяемые для последовательностей уровней НВР, считают локальными тенденциями $N\tau \in N\mathfrak{Z}$ [128].

Содержательно элементарная нечеткая тенденция моделирует изменения между двумя соседними значениями нечеткого временного ряда $\tilde{x}_i, \tilde{x}_{i+1}$ и может быть сопоставлена с моментом времени t_i .

Локальная нечеткая тенденция определяется между двумя выбранными значениями нечеткого временного ряда $\tilde{x}_i$ и $\tilde{x}_j$, при $i < j$ может быть сопоставлена с моментом времени t_i . Любая локальная НТ может быть выражена последовательностью элементарных нечетких тенденций $N\tau = \{\tau_1, \tau_2, \dots, \tau_m\}$ при $m=j-i, m < n$ [70].

Общая нечеткая тенденция характеризует поведение всего НВР и представима в виде последовательности локальных, а значит и элементарных нечетких тенденций $G\tau = \{\tau_1, \tau_2, \dots, \tau_m\}$ при $m=n$.

Последовательность элементарных тенденций, образующая локальную или общую нечеткую тенденций НВР, может быть представлена моделью нечеткой тенденции в результате применения «вычислительной» операции *Union* ACL-шкалы и операции объединения элементарных тенденций.

Таким образом, для любого НВР может быть построен временной ряд элементарных НТ и временной ряд локальных НТ [131].

Исходный нечеткий временной ряд может быть представлен в виде нечеткого временного ряда элементарной тенденции $\tau = \{t_i, \tau_i\}, i \in [1, n]$, нечеткого временного ряда локальных тенденций $N\tau = \{t_j, N\tau_j\}, j \in [1, n_1], n_1 < n$.

Таким образом, анализ временных рядов на основе нечеткой модели учитывает следующие особенности ВР:

1. Модели элементарной, локальной, общей НТ нечеткого временного ряда обладают общей структурой.
2. Локальная и общая нечеткие тенденции НВР могут быть выражены через временной ряд элементарных НТ.
3. Временной ряд элементарных НТ является инвариантным средством лингвистического представления поведения любого НВР.
4. Представление временных рядов в виде нечетких временных рядов и временных рядов элементарных нечетких тенденций позволяет учитывать при их анализе дополнительные знания в виде семантики прикладной области за счет использования контекстно-зависимых нечетких меток.

1.8 Мера ассоциации локальных трендов ВР - угловой коэффициент скользящей аппроксимации

Задача поиска сходства метрик проектной деятельности, представленных ВР, нуждается в специальных средствах выражения сходства. Авторы [4] предложили специальную меру ассоциации.

Временной ряд (y, t) или y это последовательность (y_i, t_i) , $i \in I = (1, \dots, n)$, где y_i и t_i это вещественные значения временного ряда и моментов времени, соответственно, так что $t_i < t_{i+1}$ для всех $i = 1, \dots, n-1$. Окно W_i длины $k > 1$ это последовательность индексов $W_i = (i, i+1, \dots, i+k-1)$, $i \in \{1, \dots, n-k+1\}$. Окно длины k обозначается как $y_{W_i} = (y_i, y_{i+1}, \dots, y_{i+k-1})$. Последовательность $J = (W_1, W_2, \dots, W_{n-k+1})$ всех окон длины k , $(1 < k \leq n)$, называется скользящим окном.

Линейная функция $f_i = a_i t + b_i$ с параметрами $\{a_i, b_i\}$ минимизирующая критерий

$$Q(f_i, y_{W_i}) = \sum_{j=i}^{i+k-1} (f_i(t_j) - y_j)^2 = \sum_{j=i}^{i+k-1} (a_i t_j + b_i - y_j)^2,$$

называется линейной регрессией или линейной аппроксимацией y_{W_i} . Методы вычисления оптимальных значений параметров a_i, b_i хорошо известны, например, a_i может быть подсчитан следующим образом:

$$a_i = \frac{k \sum_{j=i}^{i+k-1} t_j y_j - (\sum_{j=i}^{i+k-1} t_j) \sum_{j=i}^{i+k-1} y_j}{k \sum_{j=i}^{i+k-1} t_j^2 - (\sum_{j=i}^{i+k-1} t_j)^2}.$$

Преобразование $MAP_k(y, t) = a$, где $a = (a_1, \dots, a_{n-k+1})$, суть последовательность угловых коэффициентов a_i линейных аппроксимаций временного ряда в скользящем окне длины k , называется преобразованием скользящих аппроксимаций (САП трансформ), а значения $a = (a_1, \dots, a_{n-k+1})$, называются локальными трендами [4]. Часто в практических приложениях временные ряды имеют постоянный шаг $h = t_{i+1} - t_i$, $i = 1, \dots, n-1$. В [5] предложен способ вычисления САП трансформ для таких временных рядов:

$$a_i = \frac{6 \sum_{j=0}^{k-1} (2j-k+1) y_{i+j}}{hk(k^2-1)}, i \in \{1, 2, \dots, n-k+1\}.$$

Из формулы САП трансформа получаются следующие формулы для подсчета углового коэффициента скользящих аппроксимаций для окон размером $k = 2, 3, 4$, соответственно:

$$a_i = y_{i+1} - y_i, k = 2,$$

$$a_i = (y_{i+2} - y_i)/2, k = 3,$$

$$a_i = (3y_{i+3} + y_{i+2} - y_{i+1} - 3y_i)/10, k = 4,$$

Пусть $y = (y_1, \dots, y_n)$ и $x = (x_1, \dots, x_n)$ это два временных ряда, определенные в одни и те же моменты времени, и $MAP_k(y) = (a_{y1}, \dots, a_{ym})$, $MAP_k(x) = (a_{x1}, \dots, a_{xm})$, $k \in \{2, \dots, n-1\}$, $m = n - k + 1$ суть САП преобразования этих рядов для некоторого k . Функция

$$coss_k(y, x) = \frac{\sum_{i=1}^m a_{yi} \cdot a_{xi}}{\sqrt{\sum_{i=1}^m a_{yi}^2} \cdot \sqrt{\sum_{j=1}^m a_{xj}^2}}$$

называется мерой ассоциации локальных трендов временных рядов y и x . Эта мера равна косинусу угла между векторами САП трансформы y и x :

$$coss_k(y, x) = \cos(\angle(MAP_k(y), MAP_k(x))).$$

Пусть (y, t) и (x, t) это два временных ряда и $L_1(y, t) = (p_1 y + q_1 r_1 t + s_1)$, $L_2(y, t) = (p_2 y + q_2 r_2 t + s_2)$, где $p_1, p_2, r_1, r_2 \neq 0$, – линейные преобразования временных рядов, тогда выполняется следующее свойство инвариантности этой меры:

$$coss_k(L_1(y, t), L_2(x, t)) = \text{sign}(p_1) \cdot \text{sign}(r_1) \cdot \text{sign}(p_2) \cdot \text{sign}(r_2) \cdot coss_k((y, t), (x, t)).$$

Мера $|coss|$ инвариантна относительно независимых линейных преобразований временных рядов. Из этого следует, что результат меры ассоциаций локальных трендов не зависит от таких преобразований временных рядов как смещение нулевой точки времени, нормализация, замена значений времени их индексами при постоянном шаге и др.

Наиболее полную информацию об ассоциациях между временными рядами y, x дает последовательность ассоциаций локальных трендов, посчитанных для всех размеров скользящих окон $k = 2, \dots, n$. Такая последовательность называется ассоциативной функцией, значения которой зависят от размера:

$$AF(y, x) = (coss_2(y, x), \dots, coss_n(y, x)).$$

График этой функции дает полезную информацию об ассоциациях между временными рядами. Можно рассматривать некоторое подмножество $K \subseteq \{2, \dots, n\}$ всех возможных окон и говорить о функции $AF_K(y, x)$, определенной на

множестве окон K . Среднее или максимальное значение этой функции может быть использовано как мера ассоциации между временными рядами, например:

$$AM(y, x) = \text{mean}(AF_K(y, x)) = \overline{AF_K}(y, x) = \frac{1}{|K|} \sum_{k \in K} \cos s_k(y, x).$$

Ассоциативная функция позволяет выявлять положительные и отрицательные ассоциации между временными рядами, имеющими сложную структуру. Для временных интервалов разной длины эти ассоциации могут иметь разный характер. Таким образом, данная мера имеет преимущества по сравнению с известными мерами ассоциации и сходства между временными рядами не только благодаря ее инвариантности, но также благодаря ее способности оценивать ассоциации между временными рядами для разных временных интервалов [82].

1.9 Оценки качества ВР

Качество анализа ВР выражается в надежности прогноза (модели прогнозирования) его поведения в будущем. **Качество прогнозов оценивается** путем сравнения значений зависимой переменной y и предсказанных значений для каждого значения независимой переменной t (периоды времени). **Отклонения (остатки)** являются необъяснимой частью каждого наблюдаемого значения, позволяют проверить, применима ли данная модель и те предположения, на которых она основана, а также можно ли использовать ее для того, чтобы дать оценку вероятностных отклонений прогнозов, сделанных на основе линейной модели.

1.9.1 Абсолютные отклонения прогнозов

Первая группа оценок качества базируется на расчете абсолютных ошибок и включает в себя оценки, основанные на вычислении величины e_t [123]

$$e_t = (y_t - f_t^{(m)}),$$

где y_t – измеренное значение в момент времени t , $f_t^{(m)}$ – прогнозное значение в момент времени t , полученное в результате использования исследуемой модели прогнозирования m .

Средняя абсолютная ошибка (*Mean Absolute Error, MAE*) рассчитывается по формуле:

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| = \text{mean}_{i=\overline{1,n}} |e_i|,$$

где n – размер горизонта прогнозирования (*forecast horizon*), $\text{mean}(\bullet)$ – операция нахождения среднего.

Средняя абсолютная ошибка исчисляется путем усреднения разности между прогнозной и фактической величиной. "Абсолютной" ошибкой она называется потому, что пренебрегает знаком разности (плюс или минус) между прогнозом и фактическим объемом.

Квадрат среднеквадратичной ошибки (*Mean Square Error, MSE*) рассчитывается по формуле:

$$MSE = \frac{1}{n} \sum_{i=1}^n (e_i^2) = \text{mean}_{i=\overline{1,n}} (e_i^2),$$

следовательно, средняя квадратичная ошибка (*Root Mean Square Error, RMSE*) рассчитывается согласно формуле:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i^2)} = \sqrt{\text{mean}_{i=\overline{1,n}} (e_i^2)},$$

Поэтому процент ошибки RE (*Ratio Error*) можно найти по следующей формуле:

$$RE = (RMSE / \text{Ср. арифм.}) * 100\%,$$

где Ср. арифм. – средняя арифметическая фактического ряда поступлений.

Согласно автору статьи [57] при $RE < 5\%$ метод прогнозирования считается точным.

Кроме этого $RMSE$ может носить название среднеквадратичная ошибка прогнозных значений от измеренных, среднее значение квадрата ошибки прогноза. Данные критерии ошибок прогнозирования являются одними из самых распространенных [14]. Рассмотрим недостатки этих оценок.

1. В качестве основного недостатка можно выделить зависимость критериев от шкалы измерения [21]. Следовательно, если стоит задача оценивания моделей

прогнозирования на ряде объектов с различными шкалами измерения, то данные оценки неприменимы.

2. Другим недостатком предложенных методов является влияние на оценку выбросов в данных. Так, если в данных присутствует выброс, имеющий максимальное значение, что является распространенным случаем на практике [45], то рассчитанное значение ошибки может быть сильно завышенным. Использование медианы, дает более устойчивую оценку.

3. Оценки $RMSE$, MSE являются ненадежными: применимость данных критериев к различным порциям данных может значительно влиять на результат оценки [2].

1.9.2 Процентные (относительные) отклонения прогнозирования

Процентные ошибки рассчитываются на основании значения p_t [123]

$$p_t = \frac{|e_t|}{y_t}$$

и являются также одними из самых распространенных при решении задач прогнозирования. В отечественной литературе такие оценки называются относительной погрешностью [79] [87] [103]. В группу процентных отклонений входят следующие оценки. Среднее относительное отклонение прогноза или средняя абсолютная процентная ошибка (*Mean Absolute Percentage Error, MAPE*)

$$MAPE = \frac{1}{n} \sum_{i=1}^n 100 \cdot |p_t| = \text{mean}_{i=\overline{1,n}}(100 \cdot |p_t|),$$

Средняя квадратичная процентная ошибка (*Root Mean Square Percentage Error, RMSPE*) вычисляется согласно:

$$RMSPE = \sqrt{\text{mean}_{i=\overline{1,n}}(100 \cdot |p_t|)^2},$$

В качестве недостатков критериев данной группы следует отметить следующие.

1. Возникновение ошибки деления на ноль, в случае, когда реальное значение равно нулю.

2. Несимметричность оценок – значение отклонения различно, если прогнозное значение находится выше или ниже реального [13] [25] [27].

3. Значительное влияние на оценку выбросов – если выброс принимает значение, значительно превосходящее максимальное «нормальное» значение [2].

4. В работе [26] показано, что данная ошибка является смещенной. Это может привести к неадекватной оценке моделей прогнозирования.

Симметричные отклонения прогнозирования

Критерии, входящие в данную группу, рассчитываются на основании значения s_t , вычисляемого по формуле:

$$s_t = \frac{|e_t|}{(y_t + f_t)}.$$

В группу входят следующие отклонения. Симметричная средняя абсолютная процентная ошибка (*Symmetric Mean Absolute Percentage Error*, *sMAPE*) вычисляется согласно:

$$sMAPE = \frac{1}{n} \sum_{i=1}^n 200 \cdot |s_i| = mean_{i=\overline{1,n}}(200 \cdot |s_i|).$$

Во избежание проблем связанных с делением на ноль была предложена модифицированное *sMAPE* – *Modified sMAPE*, *msMAPE*, в знаменателе которого появился дополнительный член:

$$msMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - f_i|}{(|y_i| + |f_i|)/2 + S_i},$$

$$\text{где } S_i = \frac{1}{i-1} \sum_{k=1}^{i-1} |y_k - \overline{y_{1-1}}|, \overline{y_{1-1}} = \frac{1}{i-1} \sum_{k=1}^{i-1} y_k.$$

Отметим следующие недостатки.

1. Несмотря на свое название, данная ошибка также является несимметричной [25].

2. Кроме этого, если реальное значение равно прогнозируемому, но с противоположным знаком, или оба этих значения равны нулю, то возникает ошибка деления на ноль.

3. Аналогично процентным ошибкам данные критерии подвержены влиянию выбросов.

4. При применении более сложных расчетов возникает проблема интерпретируемости вычисления оценок, что тормозит их распространение на практике [2].

1.10 Постановка задачи исследования

Необходимо построить эффективное средство управления совокупностью проектов крупной проектной организации на основе анализа и моделирования состояния метрик проектной программно-обеспеченности. Для достижения эффективности программной системы управления проектами необходимо исследовать метрики проекта, погруженного в современную систему контроля версий. Цель исследования декомпозируется на следующие задачи:

1. Выполнить сравнительный анализ методов и средства построения различных видов систем контроля версий, средств управления проектами.
2. Выполнить научный обзор методов анализа временных рядов.
3. Разработать формализованную модель состояния проектной деятельности на основе метрик проектов, загруженных в систему контроля версий проектного репозитория.
4. Разработать средство анализа состояния метрик сущностей СКВ проектов программно-обеспеченности как инструмент управления работами организации.
5. Разработать модели и алгоритмы прогнозирования нечетких временных рядов метрических характеристик проекта.
6. Разработать модели и алгоритмы прогнозирования временных рядов метрик проекта на основании мер сходства.
7. Разработать комплекс программ моделирования как подсистему автоматизации СКВ.

2 МОДЕЛЬ И АНАЛИЗ МЕТРИК ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ

2.1 Модель проектной деятельности на основе метрик систем контроля версий

Чтобы использовать в процессе управления информацию о ходе проекта, можно воспользоваться метриками о возникновении ошибок, о внесении улучшений и реализации новых функций за определенный период в процессе разработки программного обеспечения проекта. Далее необходимо анализировать полученные временные ряды (ВР) метрик с помощью методов интеллектуального анализа, получить прогноз развития проекта (рис.2.1.).

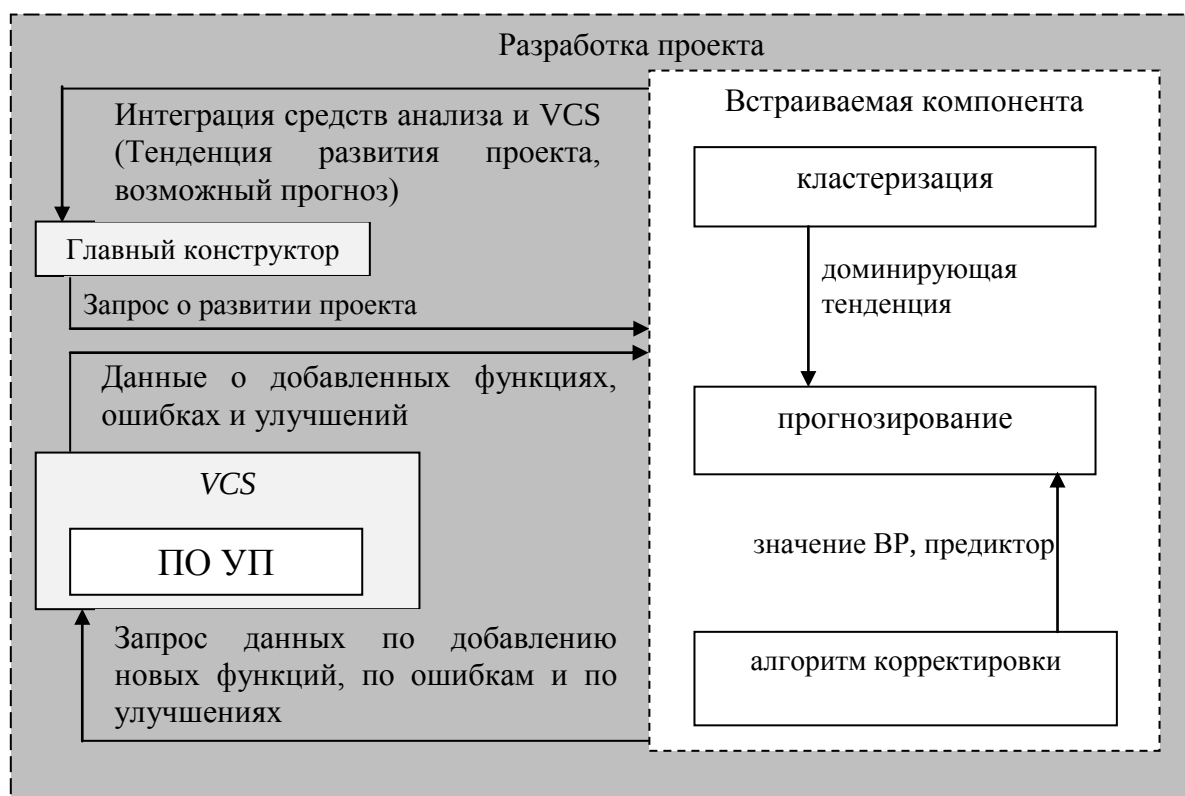


Рисунок 2.1 – Архитектура СКВ с компонентой анализа метрик

В системе контроля версий накоплено большое количество данных в виде определенных сущностей по различным проектам, атрибуты которых принимают значения разных типов: число, дата/время и др. Чтобы выявить ход работы над проектом и предсказать его будущее, проинтегрировав систему контроля версий с сервисом управления проектами, предлагается использовать следующие метрики

проекта, которые показывают количество изменений, количество стабильных выпусков проекта, число ошибок из общего числа изменений, количество улучшений из числа изменений, количество введенных новых функций за заданный период.

В процессе реализации проекта от момента утверждения стабильной версии до выпуска нового релиза ставится определенное количество задач по введению новых функций системы, исправлению соответствующих ошибок по вводу данных функций и добавлению улучшений (исправленных ошибок). Воспользовавшись знаниями эксперта и зная, что существуют пары метрик проекта, имеющие некоторый коэффициент сходства, можно предложить прогнозные значения возникновения, к примеру, ошибки на основе ее предыдущих изменений или изменений имеющей на нее влияние другой метрики. Метрика «Введение новых функций» оказывает влияние на характеристику «добавление улучшений». Следовательно, переменная «введение новых функций» является предиктором для прогнозирования других метрик. Аналогичная взаимосвязь существует между появлением ошибок и улучшениями. Появившиеся ошибки являются предиктором добавленных улучшений.

Упомянутые метрики представляются в виде ВР по определенным датам.

Модель анализа и управления проектами можно представить следующим образом [139]

$$\{C_t, R_t, B_t, I_t, F_t, R^{BI}, R^{IF}\},$$

где C_t – временной ряд commit,

R_t – временной ряд release,

B_t – временной ряд bug,

I_t – временной ряд improvement,

F_t – временной ряд New Feature,

R^{BI} – зависимость bug от improvement,

R^{IF} – зависимость improvement от New Feature.

Необходимо исследовать на сходство пары метрик, найти тенденцию поступления изменений в программное обеспечение. Одной из задач управления

проектом является задача прогнозирования. Прогноз с учетом основных задач ведения проекта (например, прогноз получения ошибок) показывает успешность проекта, ориентир на стабильную версию и позволит руководителю эффективно распределить задания сотрудникам. Для прогнозирования значений метрик предложено использовать методы нечеткого моделирования ВР.

2.2 Построение коэффициента сходства ВР. Выявление степени зависимости между ВР метрик проектной деятельности

Ввиду того, что поведение ВР, построенных по событиям в системе контроля версий, в ходе работы над проектом характеризуется высокой степенью неопределенности, а результаты востребованы в форме экспертных заключений, допускающих интерпретации в пределах допустимой точности, то для построения модели и анализа их поведения *предлагается* использовать степень сходства и корреляционную зависимость между ВР [111], [118].

Пусть задан дискретный ВР некоторой метрики $Y = \{t_i, x_i\}, i \in [1, n]$. Согласно базовым положениям теории нечетких временных рядов, разработанным L. Zade [54], Song, Chissom [47], любой конечный дискретный ВР – числовой, нечисловой и смешанный – может быть преобразован в нечеткий ВР $\tilde{Y} = \{t_i, \tilde{x}_i\}, i \in [1, n]$ при условии, что его множество значений $X = \{x_i\}$ будет покрыто специальными функциями (нечеткими множествами) $\tilde{x}_i \in \tilde{X}$.

Определение 1. Каждая нечеткая тенденция $\tau \in \mathfrak{Z}$ нечеткого временного ряда $\tilde{Y} = \tilde{x}, t = 1, 2, \dots, n$ может быть представлена структурной моделью в виде кортежа отношения, построенного на декартовом произведении свойств нечеткой тенденции $\tilde{V} \times \tilde{A} \times \Delta t \rightarrow \mathfrak{Z}$:

$$\tau = \langle \tilde{v}, \tilde{a}, \Delta t, \mu \rangle, \quad (1)$$

где τ – наименование нечеткой тенденции из множества $\mathfrak{Z}$, $\tau \in \mathfrak{Z}$;

$\tilde{v}$ – тип нечеткой тенденции (тип изменений) $\tilde{v} \in \tilde{V}$ выражает основные качественные зависимости временных рядов {Падение, Рост, Стабильность}.

$\tilde{a}$ – интенсивность нечеткой тенденции, $\tilde{a} \in \tilde{A}$, которое выражается лингвистически, например, значениями из множества $\{\text{Сильное}, \text{Среднее}, \text{Слабое}\}$;

Δt – продолжительность нечеткой тенденции, $\Delta t \in \Delta T$;

μ – функция принадлежности участка нечеткого временного ряда ограниченного интервалом Δt нечеткой тенденции τ [145].

На основе определения тенденции построим коэффициент сходства ВР, каждый из которых представлен тенденциями.

Для выявления зависимости двух ВР в работе [116] предлагается установить степень сходства поведения исследуемых ВР [62], регрессионную зависимость между ВР и найти корреляцию.

Чтобы установить степень сходства ВР НЭТ предлагается следующий критерий оценки соответствия поведения исследуемого ВР $X = x_t, t = 1, 2, \dots, n$ поведению ВР $Y = y_t, t = 1, 2, \dots, n$, который определяет расстояние между одновременными НЭТ:

$$\rho(\tau_1, \tau_2) = \frac{1}{n} \sum_{i=1}^n \left(\rho(\tilde{v}_i^{\tau_1}, \tilde{v}_i^{\tau_2}) \cdot \rho(\tilde{a}_i^{\tau_1}, \tilde{a}_i^{\tau_2}) \cdot \sqrt{\rho(\mu_i^{\tau_1}, \mu_i^{\tau_2})^2} \right),$$

где

$$\rho(\tilde{v}_i^{\tau_1}, \tilde{v}_i^{\tau_2}) = \begin{cases} 0, & \tilde{v}_i^{\tau_1} = \tilde{v}_i^{\tau_2} \\ 1, & \tilde{v}_i^{\tau_1} \neq \tilde{v}_i^{\tau_2}, \tilde{v}_i^{\tau_1} \neq C \\ 0,5, & \tilde{v}_i^{\tau_1} \neq \tilde{v}_i^{\tau_2}, \tilde{v}_i^{\tau_1} = C \end{cases},$$

$$\rho(\tilde{a}_i^{\tau_1}, \tilde{a}_i^{\tau_2}) = \begin{cases} 1, & \tilde{a}_i^{\tau_1} \neq \tilde{a}_i^{\tau_2} \\ 0, & \tilde{a}_i^{\tau_1} = \tilde{a}_i^{\tau_2} \end{cases},$$

$$\rho(\mu_i^{\tau_1}, \mu_i^{\tau_2}) = |\mu_i^{\tau_1} - \mu_i^{\tau_2}|,$$

$$0 \leq \rho(\tau_1, \tau_2) \leq 1,$$

$\rho(\tilde{v}_i^{\tau_1}, \tilde{v}_i^{\tau_2})$ – функция расстояния типов НЭТ;

$\rho(\tilde{a}_i^{\tau_1}, \tilde{a}_i^{\tau_2})$ – функция расстояния интенсивностей НЭТ;

$\rho(\mu_i^{\tau_1}, \mu_i^{\tau_2})$ – функция расстояния принадлежностей НЭТ.

Тенденция типа $\tilde{v} = C$ противоположной тенденции не имеет.

Введем метрику сходства, противоположную метрике расстояния $q = 1 - \rho(\tau_1, \tau_2)$, $0 \leq q \leq 1$. Чем больше $\rho(\tau_1, \tau_2)$, тем сильнее подобие между поведением НЭТ заданного ВР и ВР-предиктора [129].

Для проверки наличия линейной корреляционной (регрессионной) зависимости между ВР составим уравнение регрессии вида

$$\hat{y} = b_0 + b_1 x$$

Чтобы оценить тесноту корреляционной зависимости найдем выборочный коэффициент корреляции

$$r = b_1 \frac{s_x}{s_y},$$

который показывает, на сколько величина s_y изменится в среднем Y , когда X увеличится на одно s_x .

Если $r > 0$ ($b_1 > 0$), то корреляционная связь между ВР будет прямой, если $r < 0$ ($b_1 < 0$), – обратной. При прямой (обратной) связи увеличение одного из ВР ведет к увеличению (уменьшению) условной (групповой) средней другого [84].

Учитывая степень сходства и коэффициент корреляции, делаем вывод, что если существует значимая связь между ВР нечетких тенденций, то гипотеза прогноза базового ВР будет адекватна и для прогноза зависимого ВР.

2.3 Кластеризация временного ряда

2.3.1 FCM-алгоритм кластеризации

Для нечеткой кластеризации часто используют FCM-алгоритм (*Fuzzy Classifier Means, Fuzzy C-Means*) [126]. Цель FCM-алгоритма кластеризации – автоматическая классификация множества объектов, которые задаются векторами признаков в пространстве признаков. Другими словами, такой алгоритм определяет кластеры и соответственно классифицирует объекты. Кластеры представляются нечеткими множествами, и, кроме того, границы между кластерами также являются нечеткими [121], [140].

FCM-алгоритм кластеризации предполагает, что объекты принадлежат всем кластерам с определенной функцией принадлежности. Степень принадлежности определяется расстоянием от объекта до соответствующих кластерных центров. Данный алгоритм итерационно вычисляет центры кластеров и новые степени принадлежности объектов [113].

Для заданного множества K входных векторов x_k и N выделяемых кластеров c_j предполагается, что любой x_k принадлежит любому c_j с принадлежностью μ_{jk} интервалу $[0,1]$, где j – номер кластера, а k – номер входного вектора.

Принимаются во внимание следующие условия нормирования для μ_{jk}

$$\begin{aligned}\sum_{j=1}^N \mu_{jk} &= 1, \forall k = 1, \dots, K; \\ 0 < \sum_{k=1}^K \mu_{jk} &\leq K, \forall j = 1, \dots, N.\end{aligned}$$

Цель алгоритма – минимизация суммы всех взвешенных расстояний $\|x_k - c_j\|$:

$$\sum_{j=1}^N \sum_{k=1}^K (\mu_{jk})^q \|x_k - c_j\| \rightarrow \min,$$

где q – фиксированный параметр, задаваемый перед итерациями.

Для достижения вышеуказанной цели необходимо решить следующую систему уравнений:

$$\begin{aligned}\frac{\partial}{\partial \mu_{jk}} \left(\sum_{j=1}^N \sum_{k=1}^K (\mu_{jk})^q \|x_k - c_j\| \right) &= 0, \\ \frac{\partial}{\partial c_j} \left(\sum_{j=1}^N \sum_{k=1}^K (\mu_{jk})^q \|x_k - c_j\| \right) &= 0.\end{aligned}$$

Совместно с условиями нормирования μ_{jk} данная система дифференциальных уравнений имеет следующее решение:

$$c_j = \frac{\sum_{k=1}^K (\mu_{jk})^q \cdot x_k}{\sum_{k=1}^K (\mu_{jk})^q}$$

и

$$\mu_{jk} = \frac{1/\|x_k - c_j\|^{1/(q-1)}}{\sum_{j=1}^N \left(1/\|x_k - c_j\|^{1/(q-1)}\right)} \text{ (взвешенный центр гравитации).}$$

Алгоритм нечеткой кластеризации выполняется по шагам [91]:

Шаг 1. Инициализация.

Выбираются следующие параметры:

- необходимое количество кластеров N , $2 < N < K$;
- мера расстояний, как Евклидово расстояние;
- фиксированный параметр q (обычно 1,5);
- начальная (на нулевой итерации) матрица принадлежности $U^{(0)} = (\mu_{jk})^{(0)}$ объектов x_k с учетом заданных начальных центров кластеров c_j .

Шаг 2. Регулирование позиций центров кластеров.

На t -м итерационном шаге при известной матрице $\mu_{jk}^{(t)}$ вычисляется $c_j^{(t)}$ в соответствии с вышеприведенным решением системы дифференциальных уравнений.

Шаг 3. Корректировка значений принадлежности μ_{jk} .

Учитывая известные $c_j^{(t)}$, вычисляются $\mu_{jk}^{(t)}$, если $x_k \neq c_j$, в противном случае:

$$\mu_{lk}^{(t+1)} = \begin{cases} 1, l = j \\ 0, \text{ иначе} \end{cases}.$$

Шаг 4. Остановка алгоритма.

Алгоритм нечеткой кластеризации останавливается при выполнении следующего условия:

$$\|U^{(t+1)} - U^{(t)}\| \leq \varepsilon,$$

где $\| \quad \|$ – матричная норма (например, Евклидова норма);

ε – заранее задаваемый уровень точности.

2.3.2 Кластеризация ВР по доминирующей тенденции

Задача кластеризация НВР в терминах ЭТ основывается на результатах решения задачи сегментации. То есть исходными данными для кластерного анализа является множество элементарных тенденций, распределенных на временном отрезке [49].

Сформулируем постановку задачи кластеризации в терминах элементарных тенденций [112]. Пусть задано множество элементарных тенденций $\mathfrak{T} = \{\tau_i\}$, где $\tau_i = \langle \tilde{v}_i, \tilde{a}_i, \Delta t_i, \mu_i \rangle$.

Требуется разбить множество $\mathfrak{T}$ на подмножества – кластеры $s_1, s_2, \dots, s_k$ так, чтобы каждый кластер состоял из элементарных тенденций, близких к центру кластера, а элементарные тенденции разных кластеров существенно отличались [62].

На множестве $\mathfrak{T}$, с помощью метода FCM-кластеризации выделяем доминирующую тенденцию, которая преобладает в соответствующем ей кластере.

Если кластеры $s_1, s_2, \dots, s_k$ понимать как точки в параметрическом пространстве ЭТ, то задача кластеризации ВР в терминах ЭТ формулируется как выделение «сгущений элементарных тенденций» и разбиение множества $\mathfrak{T}$ на однородные подмножества элементарных тенденций.

Алгоритм кластеризации будет заключаться в построении функции f_{Clus} : $\mathfrak{T} \rightarrow S$, которая любому объекту $\tau \in \mathfrak{T}$ ставит в соответствие кластер $s \in S$.

Так как элементарные тенденции могут относиться к трем инвариантным по отношению к предметным областям и временным рядам базовым типам: «Рост», «Падение», «Стабильность», то целесообразно выбрать иерархический метод кластерного анализа, в процессе которого сначала строятся три кластера, а затем каждый кластер разбивается дополнительно на несколько кластеров, в зависимости от значения интенсивности тенденции $\tilde{a}_t$. Та тенденция, которая имеет высокую степень принадлежности к определенному кластеру и будет считаться доминирующей тенденцией выбранного временного ряда [125].

2.4 Прогнозирование ВР метрик проектной деятельности

Прогнозирование является одной из важных задач в области анализа баз данных временных рядов. Результаты прогнозирования временных рядов и тенденций, извлеченных из баз данных полезны на различных этапах управления проектами.

В частности прогнозирование рентабельности проекта, на основе анализа метрик событий при реализации данных проектов в СКВ организации является одной из основных частей в принятии решений по реализации и продолжении работы по проекту главным конструктором [110].

Существуют два подхода к задачам прогнозирования временных рядов применительно к анализу базы данных временных рядов. Первый основан на прогнозировании только одного временного ряда. Этот подход широко распространен. Для прогнозирования значений и тенденций отдельного временного ряда предложены и успешно применяются различные статистические методы [19], [81], [1], [73]; нечеткие [53], [43], [7], [47], [64]; комбинированные [33], [48], [50], [36]. Второй подход к прогнозированию временных рядов заключается в анализе статистических зависимостей между временными рядами в некотором массиве.

В этом случае в предположении, что изменения в одном из ВР является причиной изменения другого ВР строится парная регрессионная связь между временными рядами [117]. Однако часто в нестационарных временных рядах, имеющих выраженную тенденцию, при построении регрессии возникает проблема, связанная с ложной (паразитной) или неложной (действительной) линейной регрессионной связью между двумя переменными.

Данная проблема решается путем анализа коинтеграции между ВР. Если ВР не коинтегрированы, то оценивание связи между ВР не имеет смысла. Основные результаты в решении проблемы получены в работах Григори [17], Энгл [9], Йохансен [23], Филипсом [40]. Три основных способа проверки на коинтеграцию: тест Engle–Granger [9], The Johansen test [23], Phillips–Ouliaris cointegration test

[40]. Коинтегрированность является важным свойством многих переменных, для прогнозирования которых наиболее популярной является ЕСМ (ErrorCorrectionModel) — модель временных рядов, в которой краткосрочная динамика корректируется в зависимости от долгосрочной зависимости между переменными.

Применение указанных тестов и методов требует достаточно высокой квалификации от исследователя при анализе причинной связи между временными рядами, а при дальнейшем оценивании коэффициентов регрессии необходимо, чтобы не нарушались предположения регрессионного анализа [3]. Указанные факты являются проблемой для прикладного исследователя при прогнозировании временных рядов метрик проектной деятельности, в частности, если временные ряды имеют недостаточно большую длину, например, 20-40 значений.

Выходом из создавшейся ситуации может быть применение техник, применяемых для анализа одного временного ряда небольшой длины. К таким техникам относят нечеткие модели ВР [47], [61], [32] и модели нечетких тенденций [145], [63], [134].

Можно указать два ограничения в указанных моделях при прогнозе ВР: 1) моделируется только зависимость текущего состояния от предыдущих и не учитывается влияние основной нечеткой тенденции прогнозируемого ВР, 2) не учитывается ситуация, что нечеткие тенденции временного ряда Z могут выступать предиктором временного ряда Y .

Предлагается алгоритм [51], который устраняет указанные ограничения. Мы предлагаем учитывать явную гипотезу эксперта прикладной области, что нечеткие тенденции временного ряда Z — это предиктор нечетких тенденций временного ряда Y . Понятие «предиктор» используем для обозначения значимых, весомых признаков, обеспечивающих наиболее точный прогноз какого-либо явления. Иными словами, под предиктором будем понимать предпосылку какого-то важного события, выраженную численно, в составе соответствующего уравнения для прогноза.

Эта гипотеза легла в основу предложенного нового алгоритма для краткосрочного прогнозирования временного ряда Y с возможностью корректировки результатов прогноза на основе учета влияния его основной тенденции и основной тенденции временного ряда предиктора.

2.4.1 Прогнозирования ВР в терминах нечетких тенденций

Прогнозирование исходного ВР будем рассматривать как получение прогноза в терминах элементарной тенденции, при этом получение прогнозных значений нечетких и четких уровней ВР обеспечивается «оценочными» операциями ACL-шкалы, где ACL – специальная абсолютная лингвистическая шкала (*ACL - absolute comparative linguistic scale* [65]). Алгоритм прогноза временного ряда $Y = \{t_i, x_i\}, i = 1, 2, \dots, n$ будем учитывать при условии существования правдоподобности экспертной гипотезы, что нечеткая тенденция временного ряда $Z = \{t_i, z_i\}, i = 1, 2, \dots, k$ является предиктором временного ряда Y . В основу моделирования поведения ВР в терминах элементарной и локальной нечетких тенденций положим предположение о проявлении развития системы через зависимость нечеткой тенденции от значений нечетких тенденций в предыдущие моменты времени. Таким образом, структуру модели ВР будем выражать нечетким разностным уравнением d -го порядка:

$$\tau_i = \tilde{f}(t, \tau_{i-1}, \dots, \tau_{i-d}),$$

где d – фиксированное число, параметр модели;

$\tau_i, \tau_{n-1}, \tau_{n-d}$ – последовательность нечетких тенденций;

$\tilde{f}$ – некоторая нечеткая зависимость.

В [116] предлагается следующая модель ВР первого порядка для извлечения знаний о зависимостях параметров элементарных тенденций:

$$\begin{aligned}\tilde{x}_i &= Fuzzy(x_i), \\ \tilde{v}_i &= TTend(\tilde{x}_i, \tilde{x}_{i-1}), \\ \tilde{v}_{i+1} &= \tilde{f}_{\tilde{v}}(\tilde{v}_i), \\ \tilde{a}_i &= RTend(\tilde{a}_i, \tilde{a}_{i-1}),\end{aligned}\tag{2}$$

$$\begin{aligned}\tilde{a}_{i+1} &= \tilde{f}_{\tilde{a}}(\tilde{a}_i), \\ \tilde{x}_{i+1} &= \text{Comp}(\tilde{x}_{i+1}, \tilde{v}_{i+1}, \tilde{a}_{i+1}), \\ x_{i+1} &= \text{DeFuzzy}(\tilde{x}_{i+1}) + \varepsilon_{i+1},\end{aligned}$$

где *Fuzzy* – операция фаззификации ACL-шкалы по определению абсолютной нечеткой оценки $\tilde{x}_i$ по значению оцениваемого объекта x_i ;

TTend – операция ACL-шкалы по определению типа различия;

RTend – операция ACL-шкалы по определению интенсивности различия;

Comp – операция ACL-шкалы по вычислению новой абсолютной нечеткой оценки;

DeFuzzy – операции дефаззификации ACL-шкалы по определению значения оцениваемого объекта x_i по абсолютной нечеткой оценке $\tilde{x}_i$;

$\tilde{f}_{\tilde{v}}, \tilde{f}_{\tilde{a}}$ – нечеткие зависимости, представимые на основе композиционного правила импликации в виде $\tilde{v}(i-1) \circ R_{\tilde{v}}(i, i-1) \rightarrow \tilde{v}(i)$ и $\tilde{a}(i-1) \circ R_{\tilde{a}}(i, i-1) \rightarrow \tilde{a}(i)$;

x_{i+1}, ε_i – числовая оценка и ошибка прогнозируемого уровня исходного ВР.

Форма представления прогнозных значений определяется моделью, и в случае нечетких значений результат представляется в лингвистической форме для нечетких уровней, типов и интенсивностей элементарной тенденции, а для получения числовой оценки прогноза применяется операция дефаззификации ACL-шкалы.

2.4.2 Применение F-преобразования в задаче определения сходных временных рядов

Построение F-компонент. Возможна статистическая зависимость между значениями выбранного ВР и значениями другого ВР, ВР-предиктора из базы данных ВР. Проблему поиска ВР-предиктора естественно решить путем применения коэффициента корреляции. С целью сокращения вычислений коэффициент корреляции будет вычисляться по коэффициентам F-преобразования [35], [39].

Для получения информации о зависимости ВР предлагается использовать данные об изменениях локальных тенденций ВР. Ранее примененный подход, основанный на методе F-преобразования, позволял производить декомпозицию и эффективное сглаживание ВР [15], [100], [101]. Развитием такого подхода является применение F^1 -преобразования [41]. Причиной применения является то, что коэффициенты F^1 -преобразования характеризуют локальные тенденции ряда.

Для каждого ВР задаются функции нечеткого разбиения, строится F-преобразование нулевого порядка (F^0 -преобразование) [34]. Для построения F-преобразования первого порядка (F^1 -преобразование): $F_k^1 = c_{k,0}^1 + c_{k,1}^1(x - x_k)$, необходимо определить коэффициенты $c_{k,0}^1, c_{k,1}^1$. В формуле $k = 1, \dots, n$; n – количество компонент для ряда. Коэффициенты $c_{k,0}^1$ являются компонентами F^1 -преобразования. Коэффициенты $c_{k,1}^1$ определяются как:

$$c_{k,1}^1 = \frac{\sum_{t=x_{k-1}}^{x_{k+1}} f(t)(t-x_k)A_k(t)}{\sum_{t=x_{k-1}}^{x_{k+1}} (t-x_k)^2 A_k(t)},$$

где $f(t)$ – значения ВР в момент времени t , $A_k(t)$ – базисные функции нечеткого разбиения для F^1 -преобразования в момент времени t_k . Коэффициенты $c_{k,1}^1$ равны средневзвешенному значению тангенса угла наклона кусочно-линейного тренда ВР. Именно они будут служить основой анализа групп ВР. Сами компоненты F^1 -преобразования представляют собой векторы и описывают локальные тенденции ВР. В качестве примера построения одной из таких компонент приводим график ВР №104 из NN3 [29] и компоненту последнего отрезка разбиения F_n^1 :

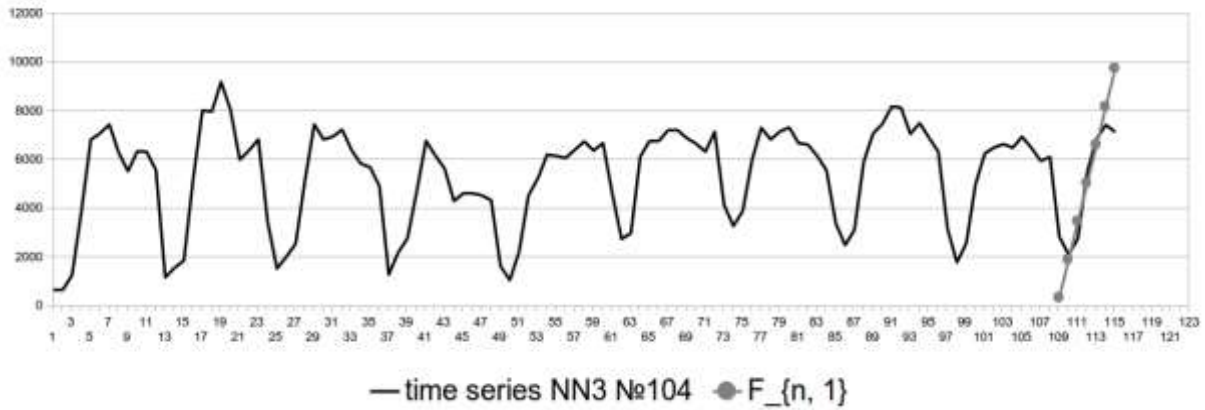


Рисунок 2.2 – ВР №104 из NN3 и компонента последнего отрезка разбиения

Формирование групп ВР [127]. Группирование ВР будем производить, анализируя коэффициенты $C_{k,1}^1$. Для каждой пары ВР вычисляется коэффициент корреляции Пирсона (линейный коэффициент корреляции):

$$r_{x'_t x''_t} = \frac{\sum_{k=1}^n (c'_{k,1} - \overline{c'_{k,1}})(c''_{k,1} - \overline{c''_{k,1}})}{\sqrt{\sum_{k=1}^n (c'_{k,1} - \overline{c'_{k,1}})^2 \sum_{k=1}^n (c''_{k,1} - \overline{c''_{k,1}})^2}}$$

где $C'_{k,1}, C''_{k,1}$ значения коэффициентов $C_{k,1}^1$ для первого (x'_t) и второго (x''_t) ВР из пары соответственно.

Важным условием для проведения анализа является равная длина ВР [20]. Перед построением компонент F-преобразования различного порядка производится нормировка значений ВР к интервалу $[0, \dots, 1]$ по модулю. Нормировка необходима для обеспечения одинаковой области значений коэффициентов $C_{k,1}^1$. При формировании групп ВР возникают следующие проблемы:

- необходимо правильно выбрать количество точек, которое покрывает базисная функция — данная величина непосредственно влияет на последующую корреляцию ВР по коэффициенту $C_{k,1}^1$. Чем большее количество точек будет покрывать базисная функция, тем более сглаженным будет ряд, и корреляция будет рассчитываться по общим тенденциям рядов;

- задание порога влияет на количество похожих по коэффициенту корреляции ВР. Важно определить, что более важно: включить в группу больше рядов, или отобрать только самые близкие по корреляции ряды.

2.4.3 Метод и алгоритм прогнозирования ВР на основе нечетких тенденций

Для прогнозирования числовых значений и нечетких тенденций состояния проекта организации в данном продукте используем метод моделирования НЭТ [115], алгоритм которого заключается в следующем [58], [59]:

1. Подсчет степени сходства поведения исследуемых ВР.
2. Составление уравнения регрессии ВР1 по ВР2. Вычисление корреляции.
3. Проверка гипотез [132]:

Гипотеза сохранения/смены тенденции. Рассматриваем тенденцию за предыдущий период. Прогнозируемое значение вычисляется по формуле:

$$\tau_{t+1} = \tau_t + \tau_p,$$

где τ_{t+1} – прогноз на следующий период времени;

τ_t – реальное значение в момент времени t ;

τ_p – значение тенденции за предыдущий период времени.

Гипотеза устойчивости/неустойчивости тенденции. Используем для прогнозирования скользящее среднее

$$\tau_{t+1} = \tau_t + G\tau_p,$$

где $G\tau_p$ – значение доминирующей нечеткой тенденции.

Рассматриваем тенденции за предыдущий выбранный период, выбираем преобладающий кластер тенденции и вычисляем прогноз по приведённой выше формуле, строим тенденцию. При одинаковых количествах вхождений тенденций используется оптимистичный прогноз, то есть выбирается наибольшее среднее значение тенденций.

Для нечеткой кластеризации тенденций используем FCM-алгоритм, основанный на минимизации целевой функции:

$$J = \sum_{j=1}^C \sum_{i=1}^N \mu_{ij}^m \|x_i - c_j\|,$$

где N – количество показателей;

C – количество кластеров;

μ_{ij} – степень принадлежности объекта i кластеру j ;

m – любое действительное число, большее 1,

x_i – i -й объект набора объектов;

c_j – j -й кластер набора кластеров;

$\|x_i - c_j\|$ – норма, характеризующая расстояние от центра кластера j до объекта i .

Объектами кластеризации являются множество элементов ВР $A = \{a_1, \dots, a_n\}$. Задачей FCM-алгоритма является разбиение этого набора на заданное количество кластеров. Вводится окончательное множество свойств (тенденций) $P = \{p_1, \dots, p_q\}$, каждое из которых количественно показывает определенный признак частей исследуемой предметной области. Где под n подразумевается общее количество объектов данных, а q – общее количество измеримых свойств.

Далее для каждого из объектов кластеризации должны быть измерены все признаки множества P в некоторой количественной шкале. Тем самым, каждому из элементов $a_i \in A$ ставится в соответствие некоторый вектор $x_i = (x_1^i, x_2^i, \dots, x_q^i)$, где x_j^i – количественное значение признака $p_j \in P$ для объекта $a_i \in A$. Все x_j^i принимают действительные значения. Эти количественные значения признаков тенденций и являются векторами признаков объектов кластеризации. Векторы значений признаков $x_i = (x_1^i, x_2^i, \dots, x_q^i)$ удобно представлять в виде матрицы данных D размерности $n \times q$, каждая строка которой равна значению вектора x_i .

Прогнозирование на заданный период на основе НЭТ. Алгоритм прогнозирования на период на основе тенденций работает следующим образом.

Эксперт задает количество рассматриваемых тенденций за предыдущий период, например, за полгода - набор тенденций A . Либо задает шаблон набора тенденций, зная предположительную тенденцию следующую за этим набором.

$$\{\tau_{t_{n-m}}, \dots, \tau_{t_{n-1}}, \tau_{t_n}\}$$

Далее идет поиск набора A тенденции во всех остальных предыдущих периодах.

$$\{\tau'_{t_{n-l-k}}, \dots, \tau'_{t_{n-l-(k-1)}}, \tau'_{t_{n-l}}\}$$

Найдя такой набор B , учитываем какая тенденция C стоит после этого найденного набора B . Строим прогноз равный тенденции C .

$$\tau_{t+1} = \tau_t + \tau'_{t_{n-l+1}}$$

Если не существует такого набора B , который совпадал бы с набором A , поиск набора повторяется, но уже ищется не полное его совпадение, а с отклонением в одну тенденцию. Так повторяем до тех пор, пока не будет найден подходящий набор тенденций B .

Выбор наилучшей гипотезы осуществляется с учетом наименьшей средней относительной ошибки (MSE).

2.5 Алгоритм коррекции гипотез прогноза зависимых временных рядов

2.5.1 Прогнозирование временного ряда посредством алгоритма адаптации

Для решения задачи прогнозирования временного ряда $Y = \{t_i, x_i\}$, ($i = 1, 2, \dots, n$) с помощью нечеткого подобия, **при условии существования правдоподобности экспертной гипотезы**, что нечеткая тенденция временного ряда $Z = \{t_i, z_i\}$, ($i = 1, 2, \dots, k$) является **предиктором временного ряда Y** используем алгоритм коррекции гипотез прогноза [117], [137]. **Алгоритм 1** включает три фазы.

На первой фазе прогнозируются нечеткие элементарные тенденции временного ряда Y , согласно (2):

$$\tau_{t+1}^Y = f(\tau_t^Y),$$

здесь τ_{t+1}^Y – прогнозная нечеткая элементарная тенденция временного ряда Y ;

τ_{t+1}^Y – текущая нечеткая элементарная тенденция временного ряда Y ;

f – зависимость в нечетких элементарных тенденциях временного ряда Y .

На второй фазе прогнозная нечеткая элементарная тенденция временного ряда Y корректируется с учетом компонент основных тенденций исследуемого временного ряда $G\tau_Y$ и временного ряда предиктора $G\tau_Z$ соответственно:

$$\hat{\tau}_{t+1}^Y = r(\tau_{t+1}^Y, G\tau_Y, G\tau_Z),$$

где

τ_{t+1}^Y – прогнозная нечеткая элементарная тенденция временного ряда Y ;

$\hat{\tau}_{t+1}^Y$ – прогнозная нечеткая элементарная тенденция временного ряда Y после корректировки;

$G\tau_Y$ – основная нечеткая тенденция временного ряда Y ;

$G\tau_Z$ – основная нечеткая тенденция временного ряда Z ;

r – правила корректировки.

Третья фаза используется для получения оценки прогнозного значения числового временного ряда Y , согласно (2).

На этой основе предлагается следующий алгоритм прогнозирования краткосрочных нечетких тенденций временных рядов.

Алгоритм 2.

Шаг 1. Преобразование числового ВР $Y = \{t_i, x_i\}$, ($i = 1, 2, \dots, n$), в нечеткий ВР $\tilde{Y} = \tilde{x}_t, \tilde{x} \in \tilde{X}, t = 1, 2, \dots, n$:

$$\tilde{x}_t = Fuzzy(x_i), x_i \in X, \tilde{x}_t \in \tilde{X}.$$

При этом интервалы, на которых определены нечеткие множества, их форма и наименование задаются пользователем исходя из особенностей предметной области.

Шаг 2. Преобразование нечеткого ВР $\tilde{Y} = \tilde{x}_t, \tilde{x} \in \tilde{X}, t = 1, 2, \dots, n$ в нечеткий временной ряд нечетких элементарных тенденций, согласно (1):

$$\begin{aligned}\tau_t^Y &= \langle \tilde{v}_t, \tilde{a}_t, \mu_t \rangle, \\ \tilde{v}_t &= TTend(\tilde{x}_t, \tilde{x}_{t+1}), \\ \tilde{a}_t &= RTend(\tilde{x}_t, \tilde{x}_{t+1}), \\ \mu_t &= \min(\mu(\tilde{x}_t), \mu(\tilde{x}_{t+1})).\end{aligned}$$

Предварительно определим множество наименований типов НТ $\tilde{V} = \{\text{Падение}, \text{Рост}, \text{Стабильность}\}$, и множество наименований интенсивностей НТ $\tilde{A} = \{\text{Сильное}, \text{Среднее}, \text{Слабое}\}$.

Шаг 3. Построение модели изменения компонент нечетких элементарных тенденций временного ряда Y и их прогнозирование на один период, используя (2):

$$\begin{aligned}\tilde{v}_{t+1} &= \tilde{v}_t \times \tilde{v}_{t-1} \times \dots \times \tilde{v}_{t-p} \circ R_{\tilde{v}}(t, t-p), \\ \tilde{a}_{t+1} &= \tilde{a}_t \times \tilde{a}_{t-1} \times \dots \times \tilde{a}_{t-p} \circ R_{\tilde{a}}(t, t-p).\end{aligned}$$

Шаг 4. Прогнозирование числового временного ряда Y с предварительной дефазификацией по формуле (2) компонент нечеткой тенденции $\tau_{t+1}^Y = \langle \tilde{v}_{t+1}, \tilde{a}_{t+1}, \mu_{t+1} \rangle$:

$$x_{t+1} = x_t + v_{t+1} \cdot a_{t+1}.$$

Шаг 5. Применение алгоритма идентификации основной тенденции (см. фаза 3. Алгоритм 1) для временного ряда Y и вычисление ее компонент $G\tau_Y = \langle \tilde{v}_{G\tau}^Y, \tilde{a}_{G\tau}^Y, \mu_{G\tau}^Y \rangle$.

Дефазификация компонент основной нечеткой тенденции временного ряда Y по формуле (2).

Шаг 6. Применение алгоритма идентификации основной тенденции (см. фаза 3. Алгоритм 1) для временного ряда Z и вычисление ее компонент $G\tau_Z = \langle \tilde{v}_{G\tau}^Z, \tilde{a}_{G\tau}^Z, \mu_{G\tau}^Z \rangle$.

Дефазификация компонент основной нечеткой тенденции временного ряда Z по формуле (2).

Шаг 7. Корректировка прогнозной нечеткой элементарной тенденции временного ряда Y $\hat{\tau}_{t+1}^Y = r(\tau_{t+1}^Y, G\tau_Y, G\tau_Z)$:

$$\hat{\tau}_{t+1}^Y = v_{t+1} \cdot a_{t+1} + v_{G\tau}^Y \cdot a_{G\tau}^Y + v_{G\tau}^Z \cdot a_{G\tau}^Z.$$

Шаг 8. Вычисление скорректированного прогнозного значения числового временного ряда Y на один период $x'_{t+1} = x_t + \hat{t}_{t+1}^Y$.

2.5.2 Правила анализа зависимости между ВР

Описанные предыдущие решения задач анализа предлагается заложить в основу инструментария для генерации правил, которые впоследствии будут дополнительно уточняться пользователем [74]. Так, полученные результаты на основе корректировок гипотез прогноза с учетом наличия минимальной ошибки MSE предлагается использовать для генерации правил следующих типов [114]:

1. Если минимальную ошибку имеет гипотеза без корректировки, то ВР Y считать не зависимым от собственной общей тенденции значений ВР Y и от ВР Z .
2. Если ошибка минимальна у гипотезы с корректировкой по основной тенденции ВР Y , то ВР Y понимать как не зависимым от ВР Z , но зависит от общей тенденции самого ВР Y .
3. Если гипотеза откорректирована по ВР Z и ее ошибка минимальна, то ВР Y рассматривать зависимым и от основной тенденции ВР Y и от ВР Z .

Данная конструкция составления правил позволят осуществить гибкий подход к анализу событий разработки проектов организации, представленных ВР.

2.6 Коэффициент подобия как способ оценки влияния наличия нечетких тенденций шаблона на прогноз

Деятельность организации представляет собой сложный процесс, характеризующийся множеством показателей, постоянно изменяющихся во времени, значения которых, измеренные через дискретные промежутки времени, образуют ВР [130], [141], [142]. Для анализа функционирования и моделирования состояния проектов предприятия предлагается использовать извлеченных из СКВ временные ряды ошибок (снятых с процесса разработки bug), а также временные ряды поставленных задач, нововведений и исправлений.

Далее рассматривается способ решения задачи поиска наиболее схожей с экспертным шаблоном последовательности НТ с помощью коэффициента соответствия и, как следствие, эффективного моделирования и прогнозирования состояния проекта.

Учитывая поведение ВР в прошедшие промежутки времени, эксперт может оценить влияние наличия шаблона поведения НТ на прогноз. В работе предлагается вычислять степень схождения экспертного шаблона НТ с исследуемым ВР НЭТ на основе степени различия НЭТ.

Степень подобия определяет точность найденной последовательности НТ с выбранным шаблоном. Если такая последовательность найдена, эксперт делает предположение о выборе наиболее верной гипотезы.

К примеру, в практике анализа фондового рынка существуют такие понятия как «атака быков», «атака медведей», наличие которых говорит о том, что если цены актива на рынке снижаются, они получают прибыль, продавая на падении рынка («атака быков»), если цены актива повышаются, они выигрывают, покупая на росте рынка («атака медведей»). Поведение игроков рынка проявляется в характерных фрагментах, шаблонах или «фигурах» ВР.

Аналогию можно провести с событиями разработки программного обеспечения проекта. Например, если временной ряд ошибок программного обеспечения одного проекта на каком-то промежутке времени будет падать, это будет говорить об аналогичном поведении подобного проекта предприятия, несмотря на то, что на том же промежутке времени его временной ряд имел тенденцию роста. Возникает необходимость возможной корректировки прогноза на негативную, несмотря на общую НТ – Рост и предшествующую прогнозу НТ – Рост средний [138].

Таким образом, предлагается метод решения задач моделирования и прогнозирования ВР поведения объектов на основе выявления степени соответствия шаблона НЭТ, предложенного экспертом, с заданным ВР НЭТ с помощью коэффициента подобия.

В разработанной программной системе (раздел 3) реализовано моделирование и прогнозирование ВР методом НЭТ.

1. Выбор ВР для анализа.
2. Преобразование исходных ВР в нечеткие с использованием алгоритмического базиса параметрической структурно-функциональной ACL-шкалы.
3. Построение модели нечеткой тенденции.
4. Выбор заранее внесенных экспертом в базу данных шаблонов НЭТ.
5. Вычисление коэффициента подобия.
6. Выбор экспертом адекватной гипотезы для прогноза.

2.7 Выводы

1. Предложенная структура автоматизации анализа метрик проекта погруженного в СКВ, отличается наличием компоненты анализа временных рядов индикаторов состояния СКВ.

2. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечеткого моделирования, использующие системы нечетких правил для различных гипотез прогноза, является результативными. Предложены три гипотезы прогноза: гипотеза сохранения тенденции, гипотеза смены тенденции и гипотезы смены тенденции на заданный период.

3. Модель и алгоритмы прогнозирования тенденций временных рядов на основе нечетких мер сходства временных рядов, отличающиеся предложенными индикаторами сходства и процедурами их построения и внесения поправок в прогноз, является эффективными.

3 РЕАЛИЗАЦИЯ ПРОГРАММНОЙ СИСТЕМЫ АНАЛИЗА И ПРОГНОЗИРОВАНИЯ СОБЫТИЙ СКВ

3.1 Информационная структура программной системы анализа и прогнозирования метрик (ПС АПМ)

3.1.1 Моделирование процесса анализа метрик проекта в нотации IDEF0/IDEF3

Для проектирования ПС АПМ использованы языки моделирования *IDEFX*, согласно ГОСТ Р501 028-2001 «Информационные технологии поддержки жизненного цикла продукции. Методология функционального моделирования» [78].

IDEF0 используется для создания функциональной модели, представляющей структуру и функции системы, а также потоки временных рядов. На рисунке 3.1 изображена схема связей функционального блока на диаграммах *IDEF0/IDEF3*.



Рисунок 3.1 – Контекстная диаграмма «Анализ состояния метрик проектов программно-аппаратных комплексов» (нотация *IDEF0*)

Декомпозиция контекстной диаграммы (рис. 3.2) содержит основные функциональные блоки, которые используются в системе.

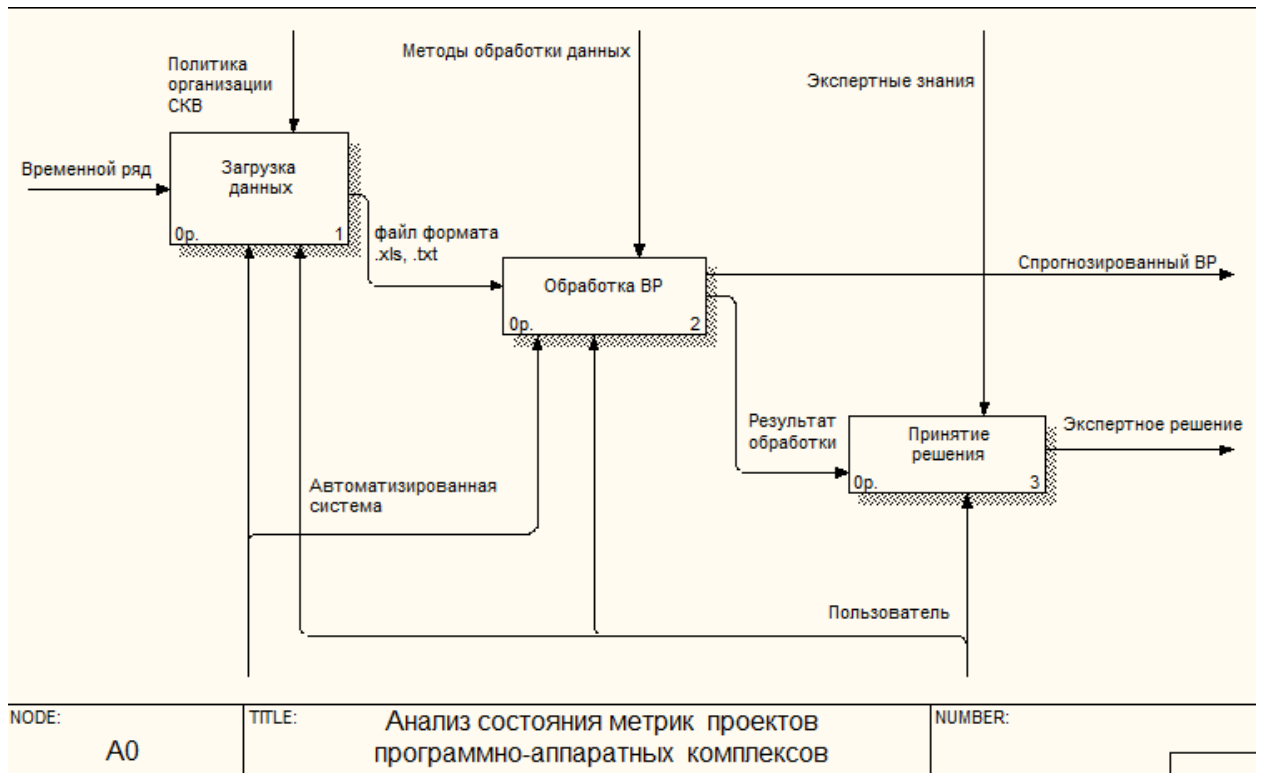


Рисунок 3.2 – Декомпозиция контекстной диаграммы (нотация *IDEF0*)

Создание модели в стандарте *IDEF3* проходит с использованием графического представления процесса, временных рядов, представленных в форматах *.xls*, *.txt* в этом процессе. С помощью *IDEF3* описывается логика выполнения работ, порядок их начала и завершения, то есть *IDEF3* предоставляет инструмент моделирования сценариев действий сотрудников организаций, отделов и т.п. Диаграмма *IDEF3* является результатом следующего шага построения модели, то есть декомпозицией функциональных блоков диаграммы *IDEF0-0*.

На рисунке 3.3 представлена декомпозиция блока «Обработка ВР» в нотации *IDEF3*.

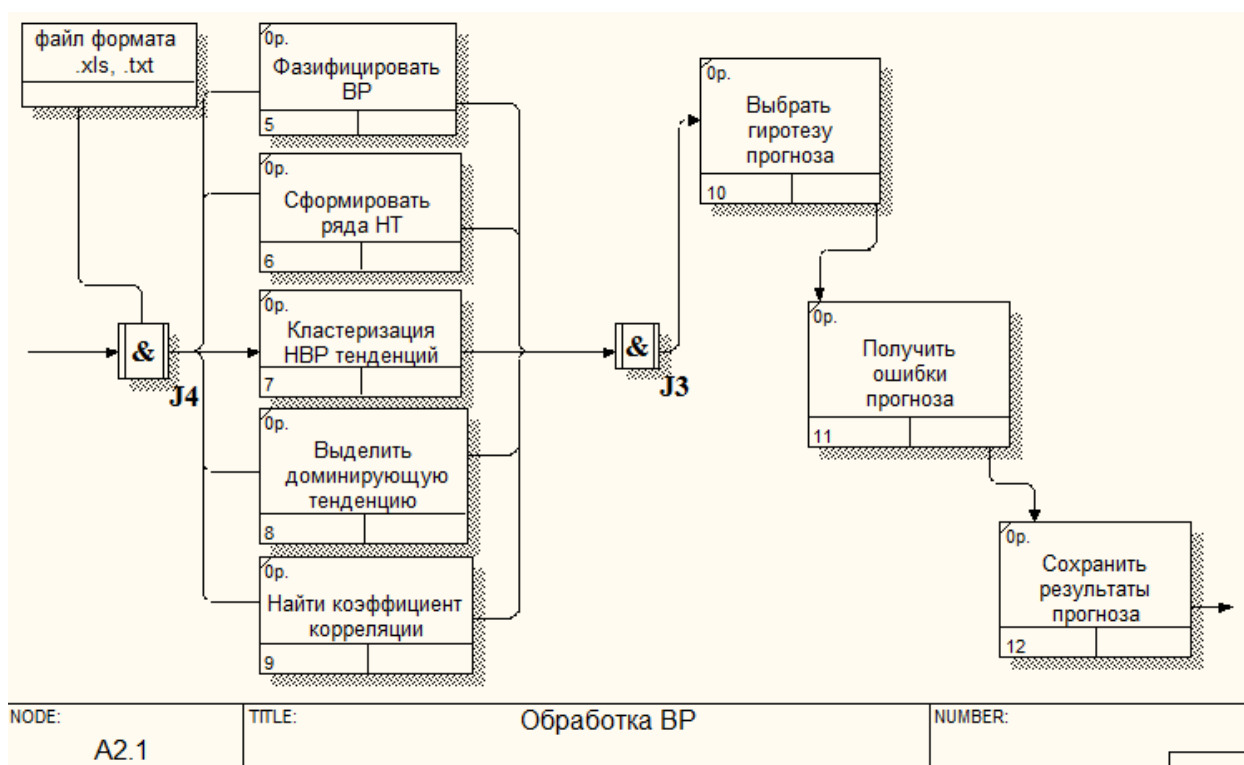


Рисунок 3.3 – Декомпозиция блока «Обработка ВР» (нотация IDEF3)

3.1.2 Структурно-функциональная схема интеллектуального хранилища

ФНПЦ АО «НПО «МАРС» многопрофильное научно-производственное предприятие, осуществляющее проектирование, разработку и поддержку научно-технической продукции. В процессе принятия решений предприятие опирается на современные средства и методы обработки информации в рамках корпоративной информационной системы предприятия.

Современный интеллектуальный репозиторий, включающий СКВ, являющийся интеллектуальным хранилищем проекта, хранит огромное количество различного рода сущностей: технические задания, проектно-конструкторская документация, инструкции, руководства, стандарты, ГОСТы, приказы, распоряжения, служебные записки, информацию о задачах, ошибках, изменениях, улучшениях и др. В рамках данного исследования ставилась задача сравнить эффективность применения различных алгоритмов анализа временных рядов с последующим прогнозированием для разных видов сущностей.

Предполагалось, что для разных видов сущностей могут оказаться наиболее эффективными разные гипотезы прогнозирования.

Программа состоит из следующих компонент: компонента описания временных рядов электронных информационных ресурсов (ЭИР), реализующая нечеткое представление ВР, метод расчета гипотез прогноза, ошибок, нормализации и денормализации; компонента кластеризации ЭИР на основе fuzzy-c-means метода («FCM-кластеризатор»).

Для работы программы требуется установка следующего программного обеспечения:

- *Microsoft Windows*;
- *Microsoft Office* версии 2007 и выше;
- *Microsoft SQL Server*.

Состав дистрибутива:

- база данных *IntellectNetArhiv* (файлы БД *MS SQL Server*) в каталоге *DB*;
- *.NET Framework 3.5 SP1*;
- *.MSChart, MSChart_VisualStudioAddOn* (для графического отображения числовых данных);
- исполняемый файл «*Analyser.exe*»;
- библиотека «*DTG.Spreadsheet.dll*»;
- документация, описание, помощь и пр. в каталоге *Doc*.

Описание используемых при индексировании таблиц БД

Словари в БД представлены двумя таблицами:

- *Templates* служит для хранения шаблонов экспертов. Хранит в себе название шаблона.
- *TemplateElements* служит для хранения информации по одной из точек шаблона. Хранит значение тенденции, значение интенсивности и функции принадлежности нечеткой тенденции.

Структуры шаблонов на рисунке 3.4.

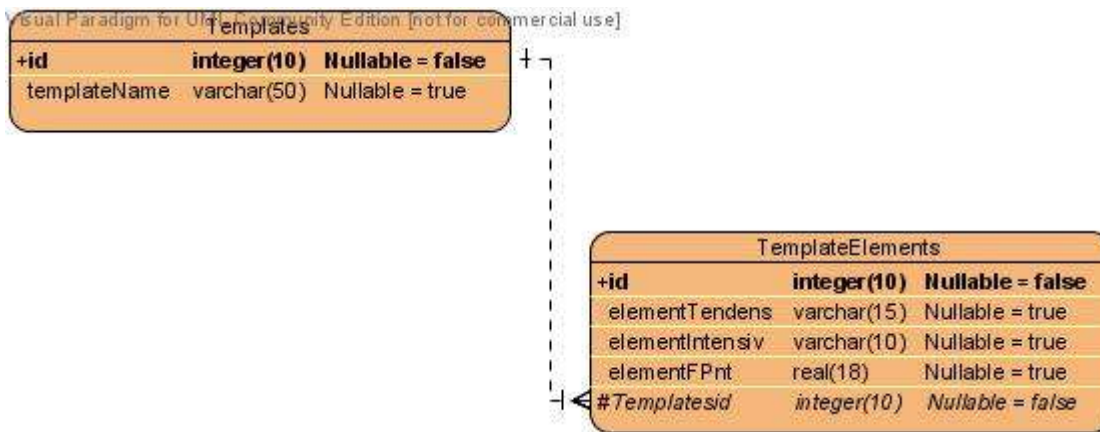


Рисунок 3.4 – Структуры шаблонов в БД

Процесс занесения данных в базу данных.

Создается макет шаблона посредством нажатия кнопки «Создать».

Далее заполняется данный шаблон элементами для анализа – тип интенсивности, тип тенденции, значение степени принадлежности.

Очищается база данных.

Для очистки шаблона из БД необходимо выбрать шаблон и нажать кнопку удалить, тем самым шаблон будет удален из БД.

Подсистема FCM-кластеризации.

В качестве входных данных «FCM-кластеризатор» использует набор ВР для ЭИР. На выходе должно быть построено дерево кластеров.

Пользователь изначально загружает файл с ВР. Затем пользователь выбирает параметры кластеризации, на основе которых происходит формирование первоначальной матрицы принадлежности ресурсов кластерам.

3.2 Обоснование инструмента реализации

Для реализации разрабатываемого решения был применен объектно-ориентированный подход. В качестве механизмов хранения информации использована реляционная система управления базами данных.

Современные открытые системы должны иметь программное обеспечение, работающее более чем на одной аппаратной платформе и операционной системе и

поддерживать изменения в аппаратной и программной части системы. Это требование выполняет кроссплатформенный язык программирования.

Так же необходимо было выбрать язык программирования, обеспечивающий хорошую поддержку взаимодействия с базами данных. Основными требованиями к среде разработки стало наличие встроенных инструментов для работы с СУБД, а также удобный редактор кода.

Средств разработки, включавших поддержку необходимых языков программирования, существует достаточно много, но не все они обладают необходимыми инструментами для работы с СУБД.

Следуя этим критериям, был выбран язык *C#* и среда разработки – *Visual Studio 2010*, имеющая удобный интерфейс взаимодействия с СУБД.

Visual Studio и язык *C#* обеспечивают полную поддержку технологии *.Net Framework*. Существует значительное количество стандартных библиотек, входящих в *.Net* для работы с базами данных. Например, *C#* был выбран благодаря поддержке такой удобной технологии как *LINQ – Language Integrated Query*, позволяющей использовать запросы, синтаксически близкие к *SQL* при обращении к данным как внутри приложения, так и к СУБД. Библиотека *LINQ* включает в себя инструмент *SQLMetal*, который позволяет автоматически генерировать классы непосредственно из поддерживаемых *.NET Framework* баз данных, что дает возможность очень быстро и просто интегрировать в код сущности базы данных.

Microsoft SQL Server — система управления реляционными базами данных, разработанная корпорацией *Microsoft*. Используется для небольших и средних по размеру баз данных, и в последние 5 лет – для крупных баз данных масштаба предприятия, конкурирует с другими СУБД в этом сегменте рынка.

3.3 Структуры данных ПС АПМ

3.3.1 Структура входных данных

Функции программы. Программа позволяет пользователю:

- загружать временные ряды из файлов формата *Excel* для прогнозирования;
- производить кластеризацию (производится автоматически при загрузке файла);
- строить прогнозные значения по трем гипотезам и производить корректировки прогнозных значений (как отдельно по ряду, так и относительно второго ряда при условии, что второй ряд является предиктором первого временного ряда);
- нормализовать ряды;
- рассчитывать ошибки прогнозных значений по каждой гипотезе;
- сохранять результаты прогнозирования в файл.

3.3.2 Структура выходных данных

Все данные полученные в результате анализа и прогнозирования сохраняются в отдельном файле в формате *.xls*. Результаты кластеризации нечеткого временного ряда и прогноза на определенный период представляются в виде диаграммы и таблиц значений.

3.4 Иерархия классов ПС АПМ

Для моделирования статического вида системы с точки зрения проектирования использована диаграмма классов (*Class diagram*). Для описываемой системы диаграмма классов представлена на рисунке 3.5.

Основным классом является класс *Series*. В главном классе создаются два экземпляра класса *Series*. Первый отвечает за работу с первым загружаемым рядом, второй экземпляр – за второй ряд, соответственно (рисунок 3.5).

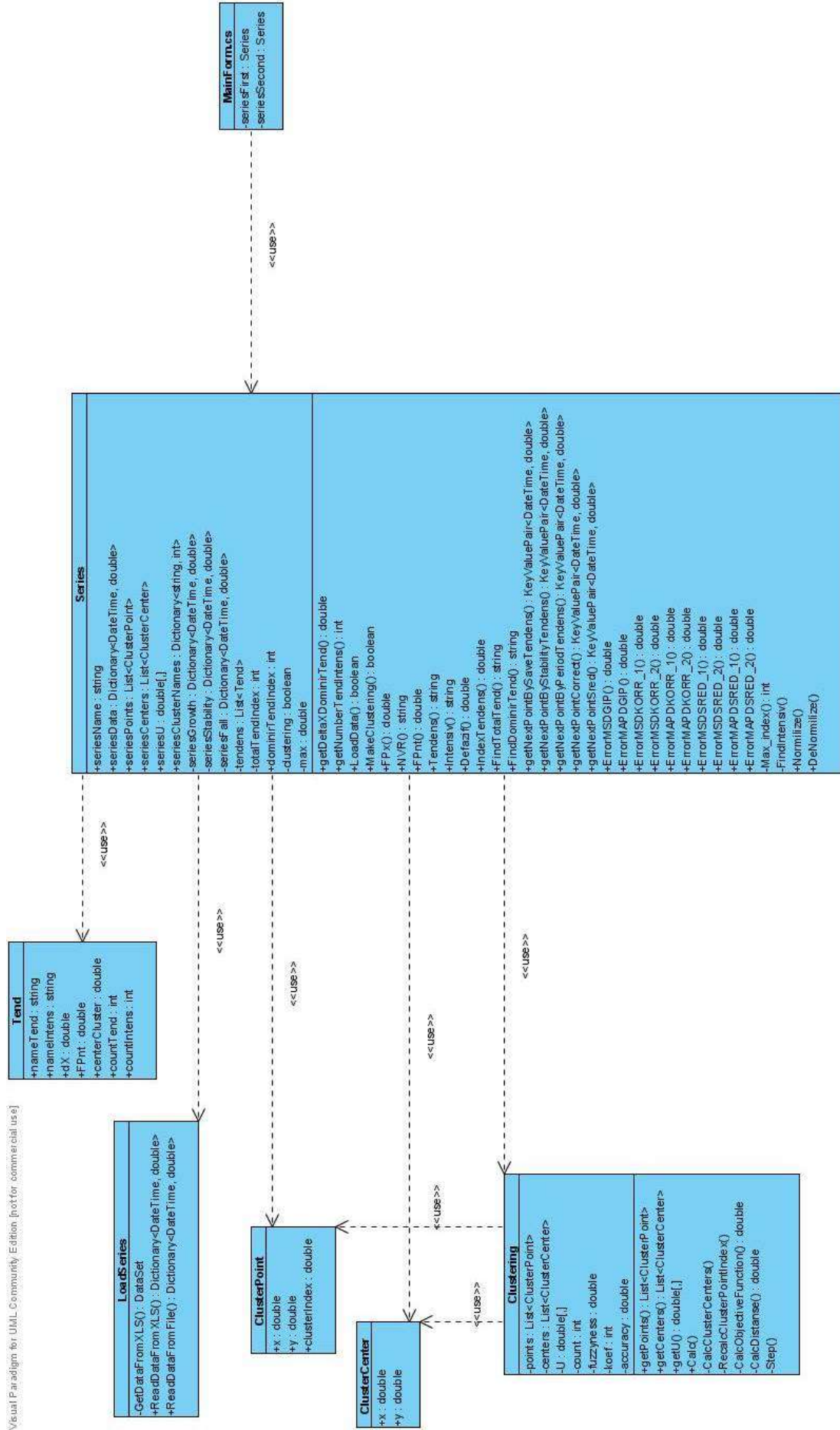


Рисунок 3.5 – Диаграмма классов

Класс LoadSeries выполняет загрузку данных из *Excel*-файла.

Данный класс содержит в себе два открытых метода:

- чтение данных из *Excel*-файла и преобразование его во временной ряд;
- чтение данных из текстового файла и преобразование его во временной ряд.

Класс *LoadSeries* включает один закрытый метод:

- формирование таблицы данных из *Excel*-файла для дальнейшего его преобразования во временной ряд.

Класс Tend. Сущностный класс, хранящий в себе информацию по статистике пар тенденция-интенсивность для расчета доминирующей тенденции и интенсивности. Имеет следующие поля:

- название тенденции;
- название интенсивности;
- сумму значений Δx всех точек данной пары тенденция-интенсивность;
- сумму значений функции принадлежности всех точек данной пары тенденция-интенсивность;
- координаты центра кластера, к которой принадлежит тенденция-интенсивность;
- счетчик встречаемости данной тенденции в ряде;
- счетчик встречаемости данной пары тенденция-интенсивность в ряде.

Класс ClusterPoint. Сущностный класс для хранения информации о точке временного ряда. Имеет следующие поля:

- координата x точки;
- координата y точки;
- номер центра кластера, к которому принадлежит точка.

Класс ClusterCenter. Сущностный класс для хранения информации о центре кластера. Имеет следующие поля:

- координата x центра;
- координата y центра.

Класс Clustering. В классе реализован механизм кластеризации. Имеет следующие поля:

- список точек кластеров;
- список центров кластеров;
- матрица нечеткого разбиения;
- количество центров кластеров;
- корректирующий значение принадлежности μ_{jk} ;
- значения порога точности для завершения расчетов.

Методы:

- метод, задающий начальные параметры расчета, формирующий матрицу нечеткого разбиения и производящий кластеризацию;
- метод для перерасчета координат кластерных центров;
- метод для расчета принадлежности точек кластера к центрам кластеров;
- метод *CalcObjectiveFunction*, определяющий погрешность перед вычислением и после него. Если разность этих погрешностей меньше порога, то вычисления заканчиваются;
- метод перерасчета матрицы нечеткого разбиения для каждой точки с последующим вызовом метода перерасчета принадлежности точек к кластерам.

Класс Series – основной класс. В нем хранится загруженный ряд, «откластеризованный» ряд, полученные тенденции и интенсивности, значение доминирующей тенденции, функции для нормализации/денормализации ряда. Данный класс имеет поля:

- название ряда;
- список исходных точек ряда в виде пар «дата – значение»;
- список точек кластеров на основе исходного ряда;
- список центров кластера;
- матрица нечеткого разбиения;
- список имен кластеров (для вывода на экран);

- список точек исходного ряда, принадлежащий к определенной тенденции;
- список класса Tend;
- список значений интенсивности в каждой точке;
- индекс общей тенденции;
- индекс доминирующей тенденции.

Методы:

- загрузка ряда;
- метод, проводящий кластеризацию ряда;
- метод, проводящий кластеризацию рядов, полученных из исходного ряда, распределением точек по тенденциям;
- метод нахождения общей тенденции;
- метод нахождения доминирующей тенденции;
- метод, возвращающий значение функции принадлежности точки;
- метод, возвращающий значение НВР точки;
- метод, возвращающий значение функции принадлежности нечеткой тенденции точки;
- метод, возвращающий значение тенденции точки;
- метод, возвращающий значение интенсивности точки;
- метод, возвращающий значение номера тенденции точки;
- метод расчета прогнозной точки ряда по гипотезе сохранения тенденции;
- метод расчета прогнозной точки ряда по гипотезе учета устойчивости тенденции;
- метод расчета прогнозной точки ряда по гипотезе на период с учетом тенденции;
- методы расчета ошибок MSE , $RMSE$ для каждой гипотезы;
- метод нормализации ряда;
- метод денормализации ряда.

3.5 Моделирование ПС АПМ

3.5.1 Диаграмма вариантов использования

Диаграмма вариантов использования (*Use case diagram*) описывает функциональное назначение системы. Она является исходным концептуальным представлением или концептуальной моделью системы в процессе ее проектирования и разработки. Для описываемой системы диаграмма вариантов использования представлена на рисунке 3.6.

Проектируемая система представляется в виде сущности – пользователя системы, взаимодействующего извне с системой с помощью вариантов использования. Пользователем могут быть конструктора, разработчики проектов. Варианты использования служат для описания сервисов, которые система предоставляет проектировщику.

Пользователь взаимодействует с вариантом использования «Провести анализ ВР». Между ним и вариантом использования существует отношение ассоциации с кратностью «1..*».

Вариант использования «Провести анализ ВР» имеет отношение расширения с вариантом использования «Выбрать ВР для анализа». При этом условием расширения является выбор пользователем временного ряда, выполнение которого служит дополнением свойств одного варианта использования благодаря наличию свойств у другого расширенного варианта использования.

Вариант использования «Провести анализ ВР» также имеет отношение обобщения с вариантом использования «Провести анализ метрик проектной деятельности». Это указывает на тот факт, что второй вариант использования может быть обобщен до первого варианта использования.



Рисунок 3.6 – Диаграмма вариантов использования

Кроме того, вариант использования «Провести анализ ВР» имеет отношения включения с вариантами использования «Осуществить ввод данных», «Представить ВР в НВР», «Откластеризовать ВР», «Осуществить прогноз НВР», «Сохранить результаты анализа». Это означает, что некоторое заданное поведение для одного варианта использования включается в качестве составного компонента в последовательность поведения другого варианта использования.

Вариант использования «Представить ВР в виде НВР» имеет отношения включения с вариантом использования «Сформировать ряд НЭТ», «Сформировать ряд интенсивностей», «Выполнить нормализацию», «Выполнить денормализацию».

Отношение расширения имеют вариант использования «Осуществить ввод данных» с вариантом использования «Получить данные из файла», «Корректировка прогноза с учетом доминирующей НЭТ второго ВР» и «Корректировка прогноза с учетом доминирующей НЭТ» с вариантом использования «Осуществить прогноз НВР». А вариант использования «Осуществить прогноз НВР» имеет в свою очередь отношение расширения с «Выбрать гипотезу прогноза» и с «Использовать шаблон эксперта».

3.5.2 Диаграмма состояний

Диаграмма состояний в *UML* описывает все возможные состояния одного экземпляра определенного класса и возможные последовательности его переходов из одного состояния в другое, то есть моделирует все изменения состояний объекта как его реакцию на внешние воздействия.

Для описываемой системы диаграмма состояний представлена на рисунке 3.7.

Каждое состояние на диаграмме деятельности соответствует выполнению некоторой элементарной операции, а переход в следующее состояние срабатывает только при завершении этой операции в предыдущем состоянии. На переходе указывается имя события и сторожевое условие, при выполнении которых объект переходит из одного состояния в другое.

3.5.3 Диаграмма последовательности

Несмотря на то, что диаграмма состояния используется для спецификации динамики поведения систем, время в явном виде на ней не присутствует. Временной же аспект поведения может иметь существенное значение при моделировании синхронных процессов, описывающих взаимодействия объектов. Для моделирования взаимодействия объектов во времени в языке *UML* используются диаграммы последовательности.

На диаграмме последовательности изображаются только те объекты, которые непосредственно участвуют во взаимодействии. Ключевым моментом для диаграмм последовательности является динамика взаимодействия объектов во времени.

Для описываемой системы диаграмма последовательности представлена на рисунке 3.8.

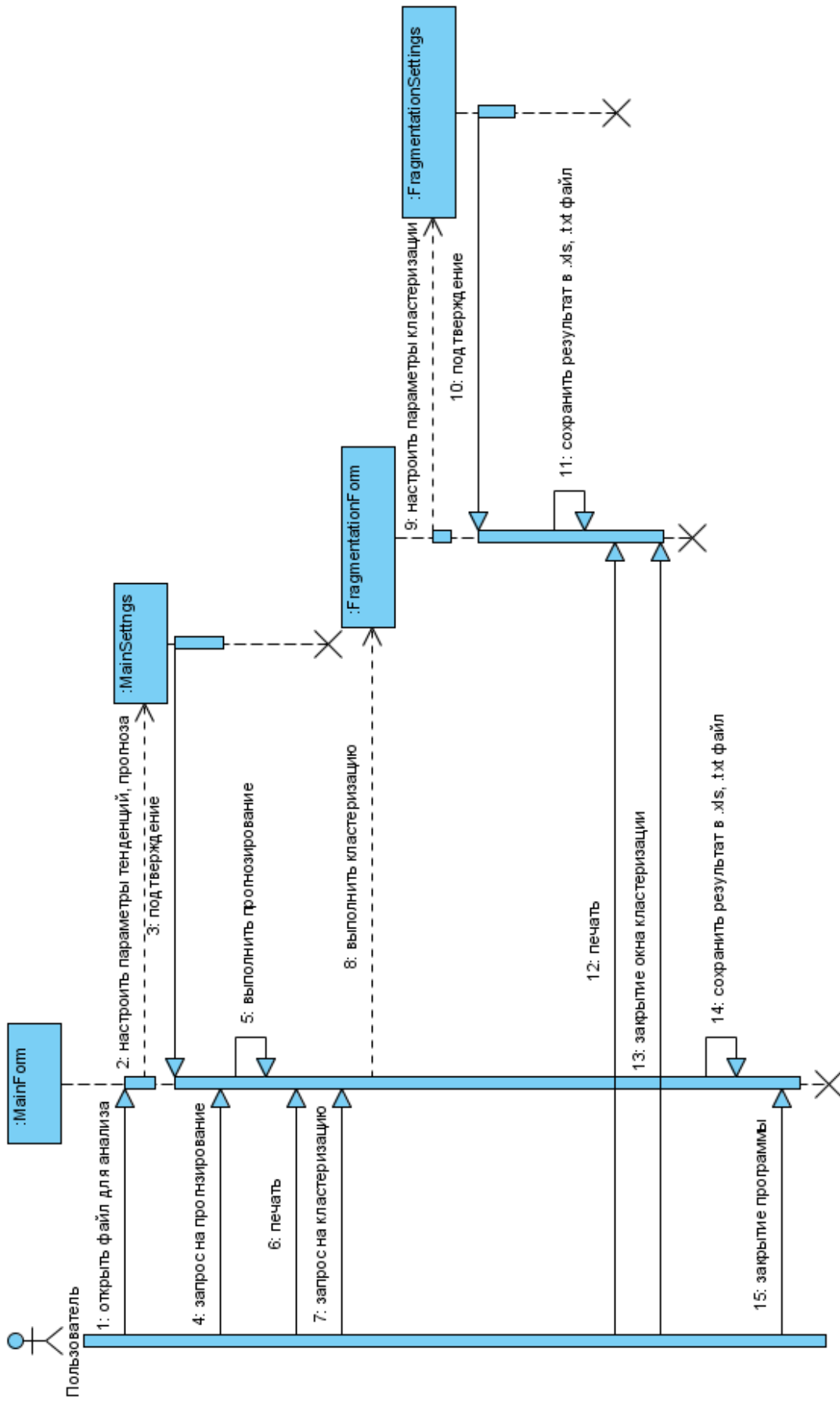


Рисунок 3.8 – Диаграмма последовательности

Объекты диаграммы последовательности появляются в системе по мере их создания пользователем. На диаграмме каждая линия жизни изображается пунктирной вертикальной линией, связанной с единственным объектом на диаграмме последовательности.

Сначала пользователь открывает форму «*MainForm*». Этот объект является постоянно активным и его фокус управления является активным на всем протяжении линии жизни. Фокус управления изображен в форме вытянутого узкого прямоугольника. Верхняя сторона – начало активности, нижняя сторона – окончание активности. Данный прямоугольник может заменить линию жизни объекта, если на всем ее протяжении он является активным.

Далее нужно добавить данные для анализа, для этого загружаем данные из файла и возвращаемся на форму «*MainForm*». После этого для настройки тенденций и значений прогноза пользователь открывает форму «*MainSettings*», на которой указывает необходимые ему параметры, после чего изменяется значение тенденций и прогноза. Потом происходит закрытие формы.

Далее пользователь запрашивает прогнозирование данных, для этого объект «*MainForm*» посылает сообщения самому себе, иницируя рефлексивные сообщения (такие сообщения изображаются прямоугольником со стрелкой, начало и конец которой совпадают). Затем пользователь запрашивает вывод на печать.

В довершении всего пользователь открывает форму кластеризации «*FragmentationForm*», где отображается график кластеризации данных. Далее для настройки количества центров кластеров пользователь открывает форму «*FragmentationSettings*», на которой указывает необходимые ему параметры, после чего изменяется результат кластеризации. Потом происходит закрытие формы. На запрос пользователя о сохранении результатов кластеризации объект «*FragmentationForm*» посылает рефлексивные сообщения. Затем пользователь запрашивает вывод на печать. Далее осуществляется закрытие формы.

На запрос пользователя о сохранении результатов анализа данных объект «*MainForm*» посылает рефлексивные сообщения. Потом происходит закрытие формы.

3.5.4 Диаграмма компонентов

Диаграмма компонентов, в отличие от ранее рассмотренных диаграмм, описывает особенности физического представления системы. Она позволяет определить архитектуру разрабатываемой системы, устанавливает зависимости между программными компонентами, в роли которых может выступать исходный и исполняемый код. Основными графическими элементами диаграммы компонентов являются компоненты, интерфейсы и зависимости между ними.

Для описываемой системы диаграмма компонентов представлена на рисунке 3.9.

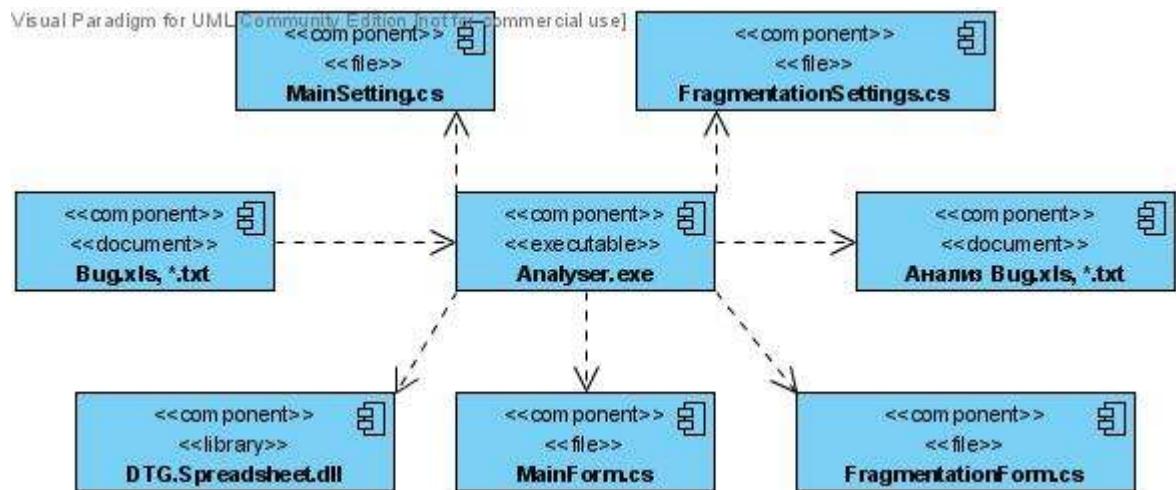


Рисунок 3.9 – Диаграмма компонентов

В процессе проектирования было выделено несколько компонентов. На диаграмме видно, что компоненты связаны между собой отношением зависимости. Отношение зависимости служит для представления связи, когда изменение одного элемента модели оказывает влияние или приводит к изменению другого элемента модели. Отношение зависимости на диаграмме компонентов изображается пунктирной линией со стрелкой, направленной от клиента (зависимого элемента) к источнику (независимому элементу).

К компонентам развертывания относится компонент – динамически подключаемая библиотека «*DTG.Spreadsheet.dll*».

К компонентам исполнения относится компонент «*Analyser.exe*».

К компонентам — рабочим продуктам относятся файлы с исходными текстами программ, а именно «*MainForm.cs*», «*MainSettings.cs*», «*FragmentationForm.cs*», «*FragmentationSettings.cs*».

Так же на диаграмме присутствуют компоненты, имеющие вторую разновидность, представляемые в форме документа. На примере анализа метрики ошибок – «*Bug.xls, *.txt*», «Анализ *Bug.xls, *.txt*».

3.6 Выводы

В третьей главе описано структурно-функциональное решение программного комплекса. Программная система, реализующая идеи нечеткого анализа метрик проектной деятельности, создана в рамках научно-исследовательского проекта по разработке автоматизированной системы «Разработка метода интеллектуального анализа многомерных нечетких временных рядов»: РФФИ № 14-01-31092, х/д НИР №Д346 от 23 сентября 2014 года УлГТУ по заказу ООО «Эверест Ресерч».

Выводы по результатам разработки программного обеспечения нечеткого подобию, представленной в Главе 3.

1. Предложенная архитектура программной системы является результативной для решения задачи анализа динамики метрик проектного репозитория.
2. Для обеспечения гибкости разрабатываемого решения был применен объектно – ориентированный подход. В качестве механизмов хранения информации выбран *Microsoft SQL Server* – система управления реляционными базами данных (СУБД) Программный комплекс разработан на языке программирования *C#* в соответствии с требованием организации-Заказчика.

4 ОЦЕНКА АДЕКВАТНОСТИ МОДЕЛЕЙ НА БАЗЕ ЭКСПЕРИМЕНТОВ И АНАЛИЗ ЭФФЕКТИВНОСТИ СРЕДСТВА АВТОМАТИЗАЦИИ ПРОЕКТНОЙ ДЕЯТЕЛЬНОСТИ

4.1 Эффекты внедрения программной системы анализа и прогнозирования метрик проектной деятельности

Успешность разрабатываемого проекта в крупной проектной организации зависит от многих факторов, в том числе от слаженной работы всей организации в целом. Обязанность руководства организации состоит в отслеживании успешности проектов, подготовки проектных документов, отчетов, вносимых изменений каждого отдела, исследовательских лабораторий принимающих участие в реализации проекта. Чтобы обнаружить проблемы, возникающие в процессе проектной деятельности, руководству необходимо постоянное информирование. Метрики проектной деятельности позволяют своевременно принимать меры по распределению ресурсов. Зная работу всего предприятия в целом и возможные слабые места, руководитель принимает определенные решения по распределению задач между группами участников проекта, которые, как показывают метрики, имеют проблемы.

Поскольку в ходе реализации проекта существуют множество событий связанных с его разработкой, главный инженер проекта (ГИП) должен учитывать каждое из них. Анализируя ход разработки посредством событий, он делает вывод о состоянии проектной деятельности. Учитывая зависимости между метриками, можно предпринять усилия, влияющие на другие метрики проекта.

При масштабах современных предприятий, занимающихся крупными проектами, отслеживание проектной деятельности занимает много времени, средств и человеческих ресурсов. Автоматизация данной деятельности ГИП обеспечит сокращение затрат, снижение нагрузки, упростит управление. Это достигается за счет снижения времени обработки данных по событиям хода проектирования,

точности анализа, получению возможного прогноза посредством гипотез и минимизации его ошибки.

4.2 План экспериментов

Для того чтобы выявить эффективность применения гипотез прогноза на основе корректировки, предложен следующий порядок действий:

- Анализируем временные ряды событий системы контроля версий.
- Выявляем пары с подобными тенденциями развития сущностей событий проекта.
- Прогнозируем по трем гипотезам, учитывая схожесть, корректируем полученный прогноз.
- Вычисляем ошибки прогнозирования. Для этого будем использовать критерии оценки точности нечеткого моделирования числовых уровней ВР в виде мер качества MSE и RMSE:

Квадрат среднеквадратичной ошибки:

$$MSE = \frac{1}{n} \sum_{i=1}^n (e_i^2) = \text{mean}_{i=\overline{1,n}}(e_i^2),$$

Средняя квадратичная ошибка:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i^2)} = \sqrt{\text{mean}_{i=\overline{1,n}}(e_i^2)},$$

- Тестируем результаты с помощью метода *ARIMA*, сравниваем ошибки.
- Эксперт на основе своих знаний и полученных данных прогноза и ошибок по каждой гипотезе – делает вывод о выборе гипотезы и об адекватности предложенной модели и метода.

4.3 Исследование временных рядов метрик на подобие и корреляцию

В описываемом эксперименте были взяты данные метрик событий проекта *MongoDB*, а именно главная его программная часть *Core Server*. Данные об ошибках, улучшениях и нововведенных функциях были взяты с ресурса *Jira.mongodb.org* [44].

MongoDB – документно-ориентированная система управления базами данных (СУБД) с открытым исходным кодом, не требующая описания схемы таблиц. Система *MongoDB* написана на языке C++.

ВР собраны за период, начиная с декабря 2014 по август 2015. В таблице 4.1 представлены результаты анализа.

Так же для оценки результативности были рассмотрены данные поступления отчетов по изделиям в архив отдела технической документации ФНПЦ АО «НПО «Марс» (Приложение Е). По данным экспертной оценки изделия можно охарактеризовать следующим образом: изделия 42 и 54 поставляются на экспорт; 275 и 220 – опытные образцы; изделия (3,41,180) и (142,14,54) являются подобными, в них достаточное количество заимствованных документов и примерно похожая конфигурация приборов [107], [108]. Результаты анализа представлены в таблице 4.2.

Таблица 4.1 – Анализ метрик событий проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
Improvement	Рост	Рост Слабый	0,61	0,5581	Средняя
New Feature	Рост	Рост Сильный			
Improvement	Рост	Рост Слабый	0,38	0,5419	Средняя
Task	Рост	Стабильность			
Bug	Падение	Рост Слабый	0,52	-0,2377	Слабая
Improvement	Рост	Рост Слабый			
Bug	Падение	Рост Слабый	0,52	0,1778	Очень слабая
New Feature	Рост	Рост Сильный			
Bug	Падение	Рост слабый	0,57	-0,4693	Умеренная
Task	Рост	Стабильность			
New Feature	Рост	Рост Средний	0,36	0,2517	Слабая
Task	Рост	Стабильность			

Таблица 4.2 – Анализ данных по поступлениям документов в архив

Коды изделий	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
14	Рост	Стабильность	0,94	0,5511	Средняя
54	Рост	Стабильность			
220	Рост	Стабильность	0,89	0,3402	Умеренная
54	Рост	Стабильность			
220	Рост	Стабильность	0,91	0,2681	Слабая
14	Рост	Стабильность			
275	Падение	Стабильность	0,81	0,2442	Слабая
142	Рост	Стабильность			
41	Рост	Стабильность	0,93	0,1749	Очень слабая
180	Рост	Стабильность			
275	Падение	Стабильность	0,82	0,0641	Очень слабая
3	Рост	Стабильность			
220	Рост	Стабильность	0,89	-0,0373	Очень слабая
41	Рост	Стабильность			
220	Рост	Стабильность	0,92	-0,0387	Очень слабая
180	Рост	Стабильность			
3	Рост	Стабильность	0,88	-0,0479	Очень слабая
180	Рост	Стабильность			
275	Падение	Стабильность	0,85	-0,0488	Очень слабая
14	Рост	Стабильность			
14	Рост	Стабильность	0,83	-0,0527	Очень слабая
142	Рост	Стабильность			
220	Рост	Стабильность	0,86	-0,0645	Очень слабая
3	Рост	Стабильность			
3	Рост	Стабильность	0,87	-0,0695	Очень слабая
41	Рост	Стабильность			
275	Падение	Стабильность	0,84	-0,0697	Очень слабая
180	Рост	Стабильность			
220	Рост	Стабильность	0,82	-0,0699	Очень слабая
142	Рост	Стабильность			
275	Падение	Стабильность	0,83	-0,0774	Очень слабая
41	Рост	Стабильность			
275	Падение	Стабильность	0,82	-0,0877	Очень слабая
54	Рост	Стабильность			
54	Рост	Стабильность	0,8	-0,1365	Очень слабая
142	Рост	Стабильность			

Степень, сила или теснота корреляционной связи определяется по величине коэффициента корреляции (r). Сила связи не зависит от ее направленности и определяется по абсолютному значению коэффициента корреляции. Для оценки

корреляционной зависимости была использована следующая градация корреляционных связей [80]:

- сильная, или тесная – при коэффициенте корреляции $|r| \geq 0,70$;
- средняя – при $0,50 \leq |r| \leq 0,69$;
- умеренная – при $0,30 \leq |r| \leq 0,49$;
- слабая – при $0,20 \leq |r| \leq 0,29$;
- очень – слабая при $|r| \leq 0,19$.

При приближении коэффициента к нулю можно говорить лишь об отсутствии линейной корреляционной зависимости (некоррелированности переменных), но не вообще об отсутствии корреляционной, а тем более, статистической зависимости.

В комплексе метрических характеристических переменных в основном между переменными значениями имеются зависимости, когда каждой величине одной переменной соответствует не какое-то определённое, а множество похожих величин другой переменной. Данная зависимость называется статистической, когда каждому значению одной переменной соответствует определённое условное распределение другой.

Появление определения статистической связи вызвано тем, что на зависимую переменную имеют влияние неконтролируемые или неучтённые обстоятельства, а также тем, что измерению значений переменных неизменно сопутствуют некоторые случайные ошибки.

Статистическая зависимость между двумя переменными, при которой каждому значению одной переменной соответствует определённое условное математическое ожидание (среднее значение) другой, называется корреляционной.

Функциональная зависимость представляет собой частный случай корреляционной. В случае функциональной зависимости при изменении значения некоторой переменной x однозначно изменяется определенное значение переменной y , при корреляционной – определённое среднее значение (математическое ожидание) y , а при статистической – определённое

распределение переменной y . Каждая корреляционная зависимость является статистической, но не каждая статистическая зависимость является корреляционной [99].

Из этого следует, что, основываясь на результатах вычисления коэффициента сходства, можно говорить о влиянии тенденций метрик СКВ (влияние «улучшений» на нововведения, «ошибок» на поставленные задачи, «улучшений» на задачи).

4.3.1 Применение F-преобразования

На ВР №104 из NN3 [29] были проведены эксперименты с применением F-преобразования в задаче определения сходства временных рядов.

Первый эксперимент выполнен на множестве ВР статистики финансовых показателей предприятий (коэффициенты: абсолютной ликвидности, автономии, срочной ликвидности, маневренности собственных оборотных средств, оборачиваемости оборотного капитала, оборачиваемости собственного капитала, текущей ликвидности, финансовой независимости в части запасов, финансовой независимости в части оборотных средств) [75]. Были отобраны ВР, имеющие одинаковую длину (в конкретном случае было выбрано 14 рядов длиной 37 точек). Заранее известно, что ряды в этом множестве обладают схожим характером поведения. Второй эксперимент был проведен на ВР соревнования NN3 (количество рядов равно 8, длина ряда – 126 точек). Количество точек, которое покрывает базисная функция, было выбрано равным 7, чтобы компоненты F-преобразования сглаживали резкие выпадения ВР.

Значения коэффициента корреляции для первого эксперимента представлены в таблице 4.3 [94]:

Таблица 4.3 – Значения коэффициента корреляции

BP 1	BP 2	Корр.	BP 1	BP 2	Корр.	BP 1	BP 2	Корр.	BP 1	BP 2	Корр.
1	2	0.97	2	13	0.99	5	6	0.85	7	14	0.96
1	3	0.91	2	14	0.96	5	7	0.84	8	9	0.94
1	4	0.73	3	4	0.89	5	8	0.78	8	10	0.9
1	5	0.84	3	5	0.96	5	9	0.92	8	11	0.97
1	6	0.62	3	6	0.8	5	10	0.92	8	12	0.95
1	7	0.95	3	7	0.91	5	11	0.8	8	13	0.97
1	8	0.94	3	8	0.85	5	12	0.78	8	14	0.99
1	9	0.95	3	9	0.96	5	13	0.87	9	10	0.99
1	10	0.94	3	10	0.97	5	14	0.77	9	11	0.94
1	11	0.93	3	11	0.87	6	7	0.55	9	12	0.92
1	12	0.95	3	12	0.87	6	8	0.45	9	13	0.98
1	13	0.97	3	13	0.92	6	9	0.72	9	14	0.93
1	14	0.94	3	14	0.84	6	10	0.77	10	11	0.92
2	3	0.93	4	5	0.92	6	11	0.51	10	12	0.9
2	4	0.74	4	6	0.95	6	12	0.48	10	13	0.97
2	5	0.86	4	7	0.68	6	13	0.63	10	14	0.9
2	6	0.63	4	8	0.59	6	14	0.45	11	12	0.97
2	7	0.97	4	9	0.81	7	8	0.96	11	13	0.97
2	8	0.96	4	10	0.85	7	9	0.95	11	14	0.97
2	9	0.98	4	11	0.63	7	10	0.93	12	13	0.95
2	10	0.97	4	12	0.61	7	11	0.98	12	14	0.96
2	11	0.96	4	13	0.73	7	12	0.98	12	15	0.54
2	12	0.96	4	14	0.57	7	13	0.97	13	14	0.97

Если выделить группы временных рядов при коэффициенте корреляции >0.9 , то получим следующий результат (табл. 4.4):

Таблица 4.4 – Группы ВР (коэффициент корреляции >0.9)

Ряд	Количество похожих рядов
1	10
2	10
3	7
4	2
5	4
6	1
7	10
8	8
9	9
10	8
11	9
12	8
13	10
14	8

Один из вариантов формирования группы по данным результатам - выделить ряды, имеющие максимально количество похожих и включить в список все участвующие ряды [92]. Например, можно выделить следующую группу рядов: 1,2,3,7,8,9,10,11,12,13,14.

Рассмотрим процедуру моделирования и краткосрочного прогнозирования ВР методом НЭТ [102], [127].

1. Выбор ВР для анализа.
2. Преобразования в нечетких исходных ВР с использованием ACL-шкалы.
3. Построение модели нечеткой тенденции.

4. Просмотр результатов в графической, числовой, лингвистической формах. Визуальный анализ (рис. 4.1) показывает хорошее качество соответствия проведенного анализа динамике исследуемого ВР.

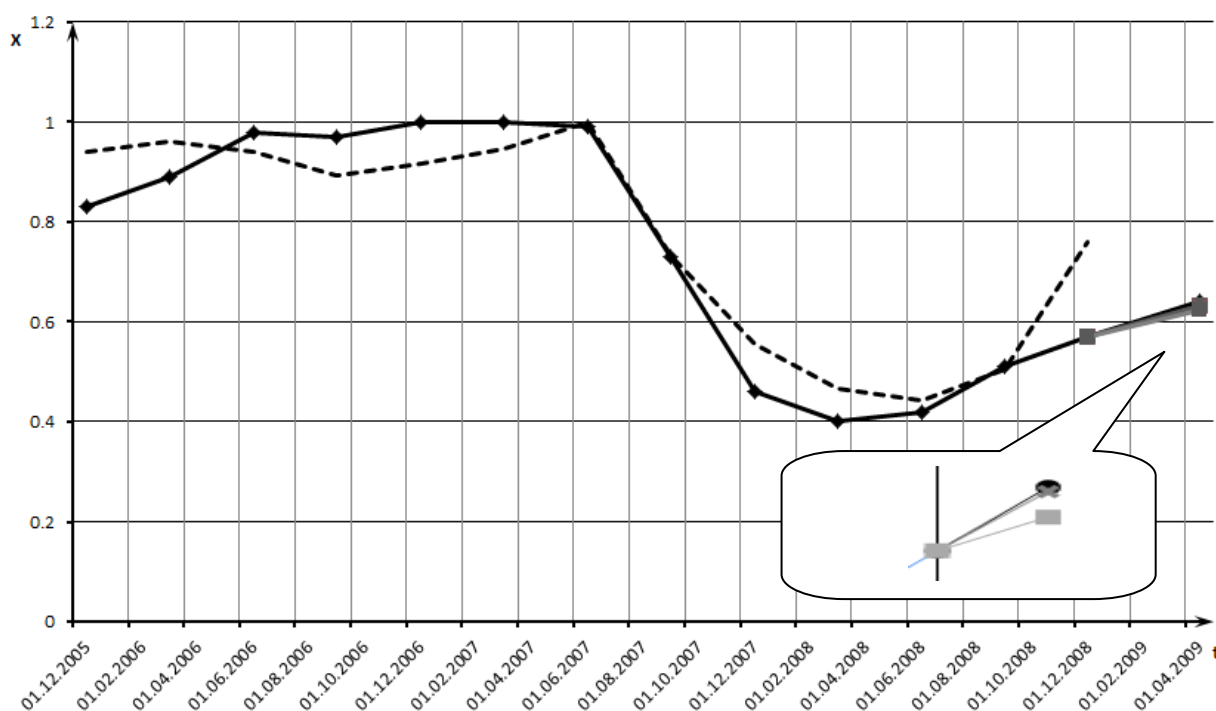


Рисунок 4.1 – Анализ состояния производственного предприятия по его ВР

На рисунке 4.1 представлен вариант прогнозирования с использованием гипотезы, где ось x – значение ВР, ось t – время; « $\blacklozenge$ » обозначает ВР Y ; «----» обозначает ВР Z ; « $\blacklozenge$ » обозначает прогноз без учета корректировки; « $\blacklozenge$ » обозначает прогноз τ_{t+1}^Y после корректировки ВР Y по основной тенденции ВР Y ; « $\blacksquare$ » обозначает прогноз $\hat{\tau}_{t+1}^Y$ после корректировки ВР Y по основной тенденции ВР Z .

Результаты анализа НВР представлены в таблице 4.5. В таблице 4.6 представлены результаты прогноза с учетом корректировок.

Таблица 4.5 – Результаты анализа

Ряд	Общая тенденция	Корреляция	Гипотеза 1
НВР1	Рост	0,9995 (сильная)	Рост слабый
НВР2	Падение		

Таблица 4.6 – Результаты прогноза

Гипотеза	Прогноз ВР			
	τ_t	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Сохранение тенденции	0,573	0,641	0,633	0,625

Поскольку корреляция сильная, делаем вывод, что ряды взаимно зависимы. Так как Y сильно зависит от Z , то и прогнозирование будет с учетом тенденции Z . В таблице 4.7 приведены оценки прогноза НВР2:

Таблица 4.7 – Критерии оценки прогноза

Оценка прогноза	Прогноз	
	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
MSE	0,006	0,0033
$RMSE$	0,0774	0,0574

Полученные результаты для данного примера свидетельствуют об адекватности и эффективности подобранной гипотезы для решения задачи прогноза и НТ ВР, содержащего данные финансовых коэффициентов, рассчитанных на базе бухгалтерских данных о деятельности организации (12 временных точек).

4.3.2 Сравнение методов преобразования с методами углового коэффициента сходства

Для сравнения результатов был проведен эксперимент на парах экономических рядов показателей финансовой организации (табл. 4.8) [38], [60], [124], [136]:

1. Коэффициент автономии / Соотношение заемных и собственных средств.
2. Коэффициент абсолютной ликвидности / Коэффициент маневренности.

3. Коэффициент текущей ликвидности / Коэффициент финансовой независимости в части запасов.
4. Коэффициент автономии / тест1.
5. Коэффициент финансовой независимости в части оборотных средств / Рентабельность всех активов по чистой прибыли.

Таблица 4.8 – Результаты сравнения

№ п/п	Мера ассоциации локальных трендов ВР у и х			Степень различия НЭТ	Степень сходства НЭТ	Корреляция	Оценка корреляции
	$\cos s_k(y, x)$			$\rho(\tau_i, \tau_s)$	$q(\tau_i, \tau_s)$		
	k=2	k=3	k=4				
1	-0,9992	-0,9996	-0,9995	0,79	0,21	-0,9995	сильная
2	0,5028	0,3455	0,4687	0,44	0,56	0,4718	слабая
3	-0,1727	-0,2398	0,0607	0,83	0,17	0,0419	нет корреляции
4	0,9267	0,9743	0,9858	0,21	0,79	0,9814	сильная
5	-0,3805	-0,3448	-0,5132	0,85	0,15	-0,4414	слабая

Результаты проведенного эксперимента представляют равенство показателей исследуемых методов, что говорит о возможности использования предлагаемых [115] способах выявления сходства между ВР.

4.4 Прогнозирование временных рядов на основе сходства поведения.

Показатели критериев качества прогноза

По данным проверки на количество одинаковых документов, участвующих в проектах, получили следующую таблицу 4.9:

Таблица 4.9 – Количество одинаковых документов входящих в архив изделия

Код изделия 1 (общее кол-во документов за равные период времени)	Код изделия 2 (общее кол-во документов за равные период времени)	Количество одинаковых документов
180 (919)	41 (549)	71
220 (985)	180 (919)	22
14 (294)	220 (985)	15
14 (294)	54 (436)	8
54 (436)	220 (985)	3
3 (2116)	180 (919)	2
54 (436)	142 (1953)	2
220 (985)	3 (2116)	2
220 (985)	41 (549)	2
220 (985)	142 (1953)	2
14 (294)	142(1953)	1

По полученным результатам анализа можно сделать следующие выводы: по статистике поступления документов в архив изделий 14 и 54 сходство и корреляция значительны, это говорит о влиянии результатов разработки изделия с кодом 54 на процесс разработки изделия с кодом 14. Аналогичные выводы по изделиям 275-142, 220-14, 220-54. Согласно таблице 4.9 корреляция между ВР поступления в архив документов изделий 180 и 220 очень слабая, несмотря на высокий показатель сходства. Так же можно добавить, что результаты разработки изделий 142,14 и 54 влияют на процесс разработки экспортных образцов изделий 275 и 220 [109]. Результаты представлены в таблице 4.10.

По результатам анализа метрик получаем: по статистике работы над проектом добавленные улучшения и введенные новые функции имеют значительные сходство и корреляцию, это говорит о влиянии результатов разработки новых функциональных возможностей на улучшения в проекте.

Аналогичны выводы по появившимся ошибкам и улучшениям. Результаты представлены в таблице 4.11.

Таблица 4.10 – Результаты прогнозирования поступления документации в архив с учетом влияния ВР-предиктора

ВР поступления документов в архив (Код изделия)	ВР-предиктор (Код изделия)	Гипотеза 1	Гипотеза 2	Гипотеза 3
14	54	Стабильность	Стабильность	Стабильность
142	275	Стабильность	Стабильность	Рост слабый
14	220	Стабильность	Стабильность	Стабильность
54	220	Стабильность	Стабильность	Падение сильное
180	41	Стабильность	Стабильность	Стабильность

Таблица 4.11 – Результаты прогнозирования метрик событий с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
Improvement	Bug	Стабильность	Рост сильный	Стабильность
Improvement	New Feature	Рост средний	Рост сильный	Рост сильный
Improvement	Task	Стабильность	Рост сильный	Стабильность
Bug	Improvement	Стабильность	Падение сильное	Стабильность
Bug	New Feature	Рост средний	Падение сильное	Рост сильный
Bug	Task	Падение среднее	Рост средний	Рост средний
New Feature	Task	Рост сильный	Рост сильный	Падение сильное

Рассмотрим подробнее результаты прогнозирования поступления документов в архив (таблица 4.12) и метрик событий проекта (таблица 4.13) с учетом влияния предиктора.

Таблица 4.12 – Прогнозы поступления документов в архив

ВР поступления документов в архив (Код изделия)	ВР- предиктор (Код изделия)	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
14	54	0	-0,00121	-0,00242	0	-0,00121	-0,00242	0,0087	0,00749	0,00628
142	275	0	-0,23285	-0,46569	17	16,76715	16,53431	28	27,76715	27,53431
14	220	0	-0,1308	-0,2616	0	-0,1308	-0,2616	1	0,8692	0,7384
54	220	16	10,11923	4,23846	16	10,11923	4,23849	8	2,11923	-3,76154
180	41	0	-0,08426	-0,16851	0	-0,08426	-0,16851	2	1,91574	1,83149

Таблица 4.13 – Значения метрик по результатам прогноза

ВР метрик	ВР-предиктор	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Improvement	Feature	25	30,71	41,99	65	70,71	81,99	26	31,71	42,99
Improvement	Task	25	30,71	32,02	65	70,71	72,02	26	31,71	33,02
Bug	Improvement	0	0	6,4	43	53,2	63,4	36	46,2	56,4
Bug	Feature	0	1,31	59,31	43	58,31	116,31	30	45,31	103,31
Bug	Task	0	0	1,7	43	53,2	58,7	36	46,2	51,7
Feature	Task	8	13,38	12,38	11	16,38	15,38	0	5,38	4,38

Графически корректировка тенденций при условии влияния временного ряда предиктора представлены на рисунках 4.2 – 4.6 (по поступлениям документации в архив) и на рисунках 4.7 – 4.12 (по метрикам событий проекта).

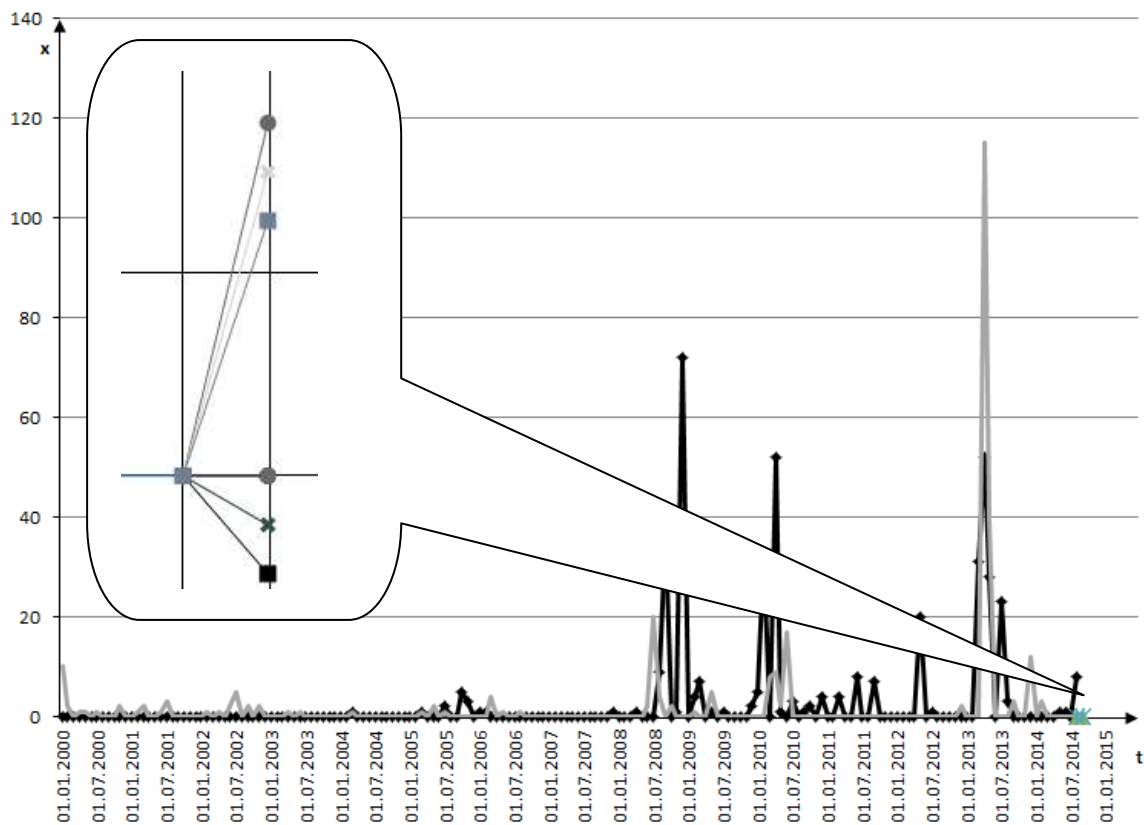


Рисунок 4.2 – Прогнозирование поступления документации изделия с кодом 14 с учетом влияния изделия с кодом 54

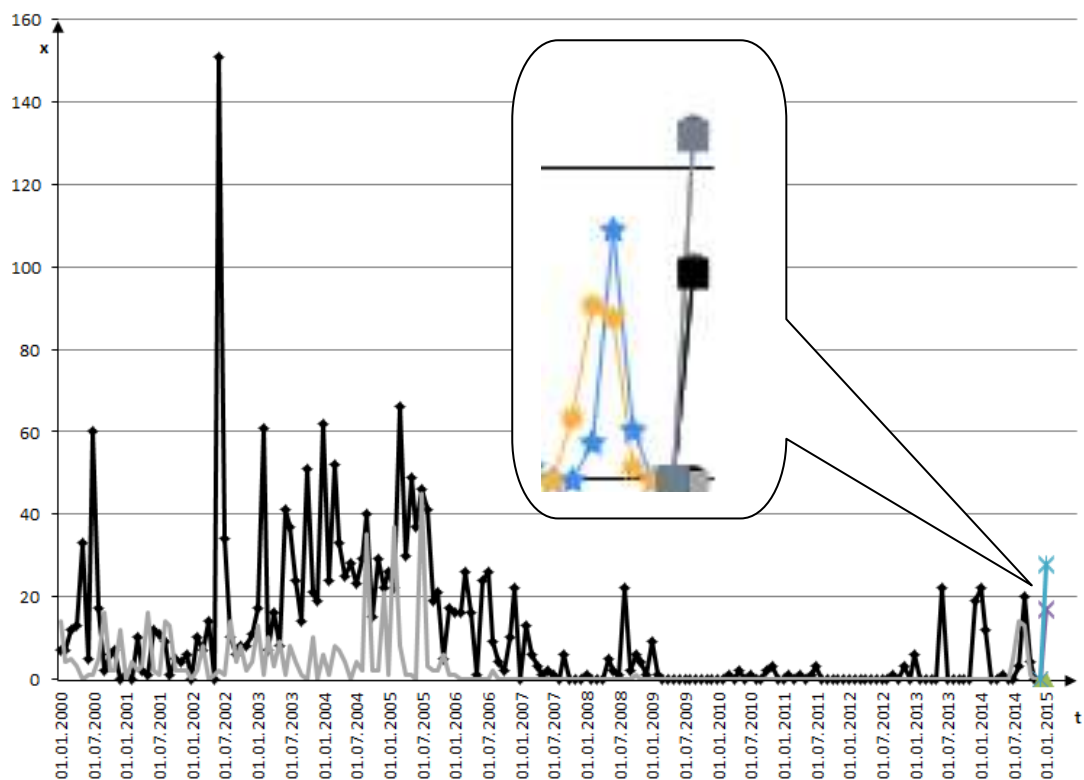


Рисунок 4.3 – Прогнозирование поступления документации изделия с кодом 142 с учетом влияния изделия с кодом 275

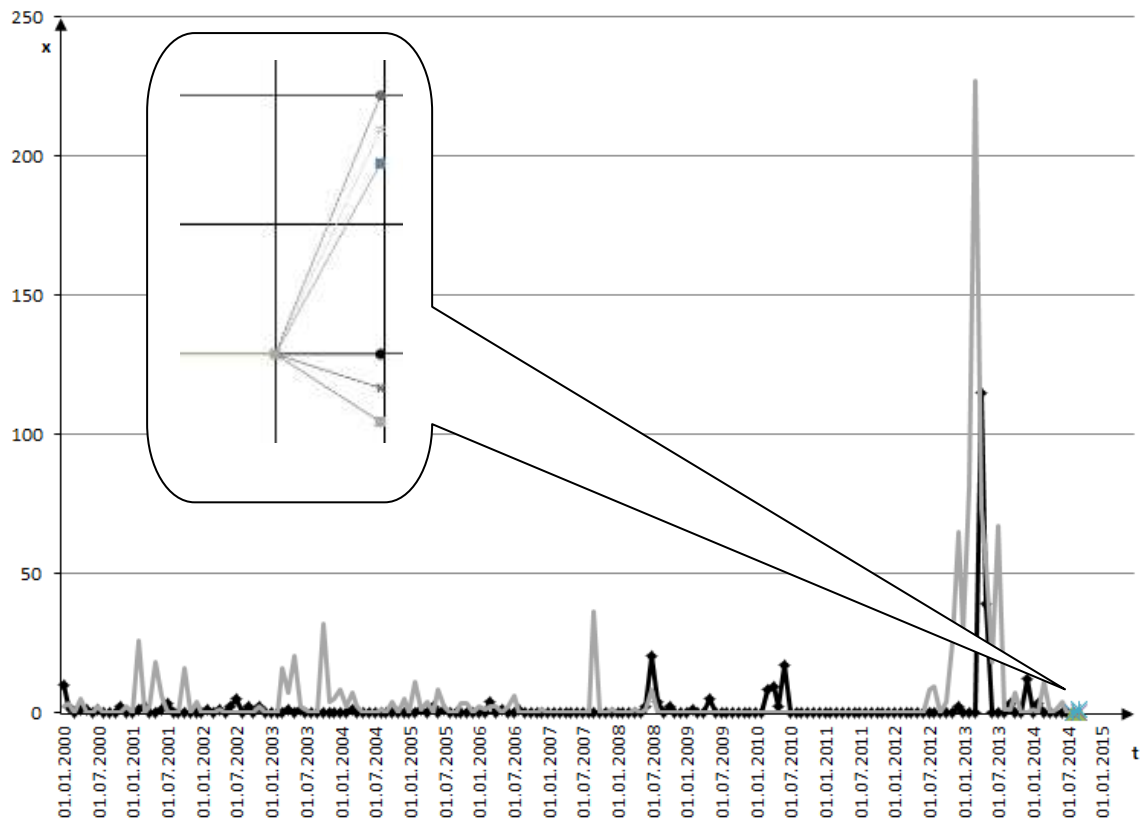


Рисунок 4.4 – Прогнозирование поступления документации изделия с кодом 14 с учетом влияния изделия с кодом 220

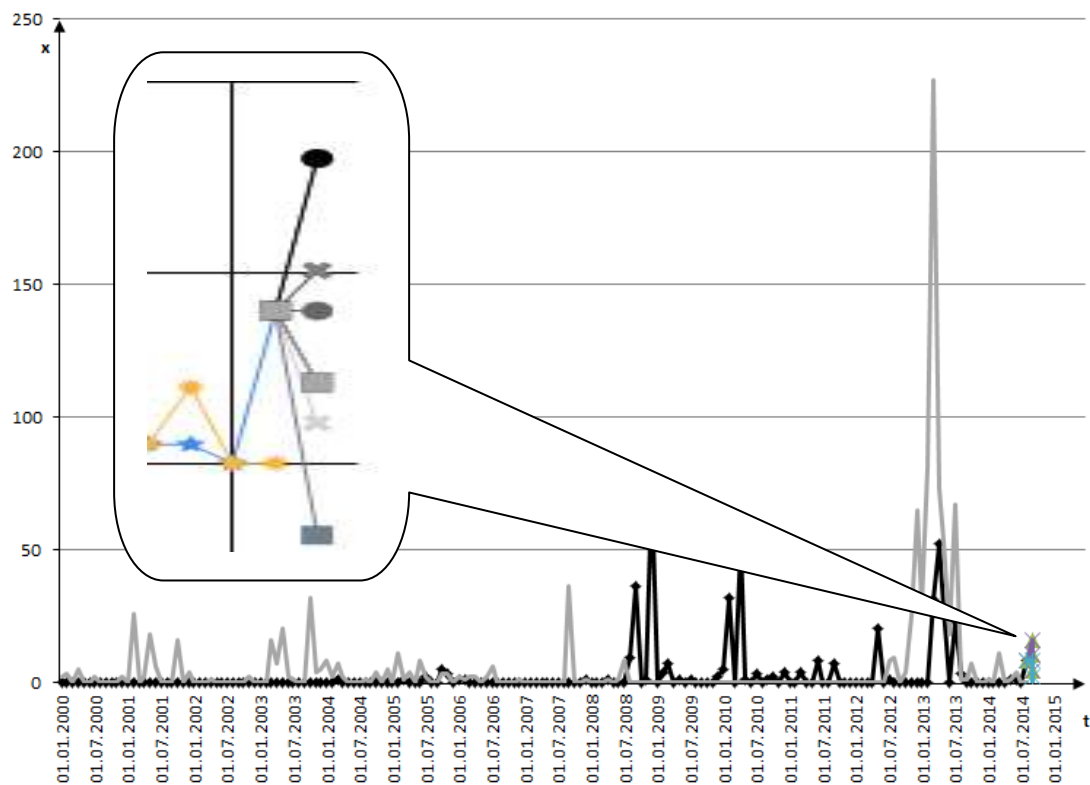


Рисунок 4.5 – Прогнозирование поступления документации изделия с кодом 54 с учетом влияния изделия с кодом 220

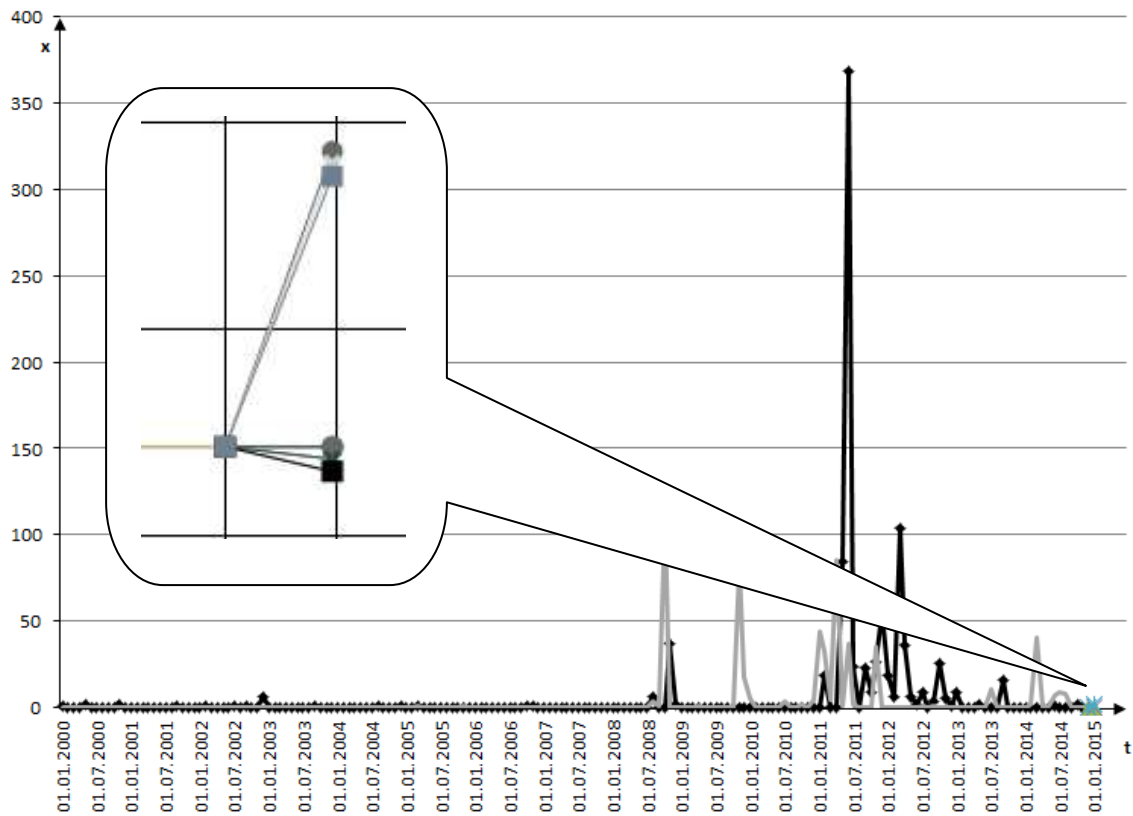


Рисунок 4.6 – Прогнозирование поступления документации изделия с кодом 180 с учетом влияния изделия с кодом 41

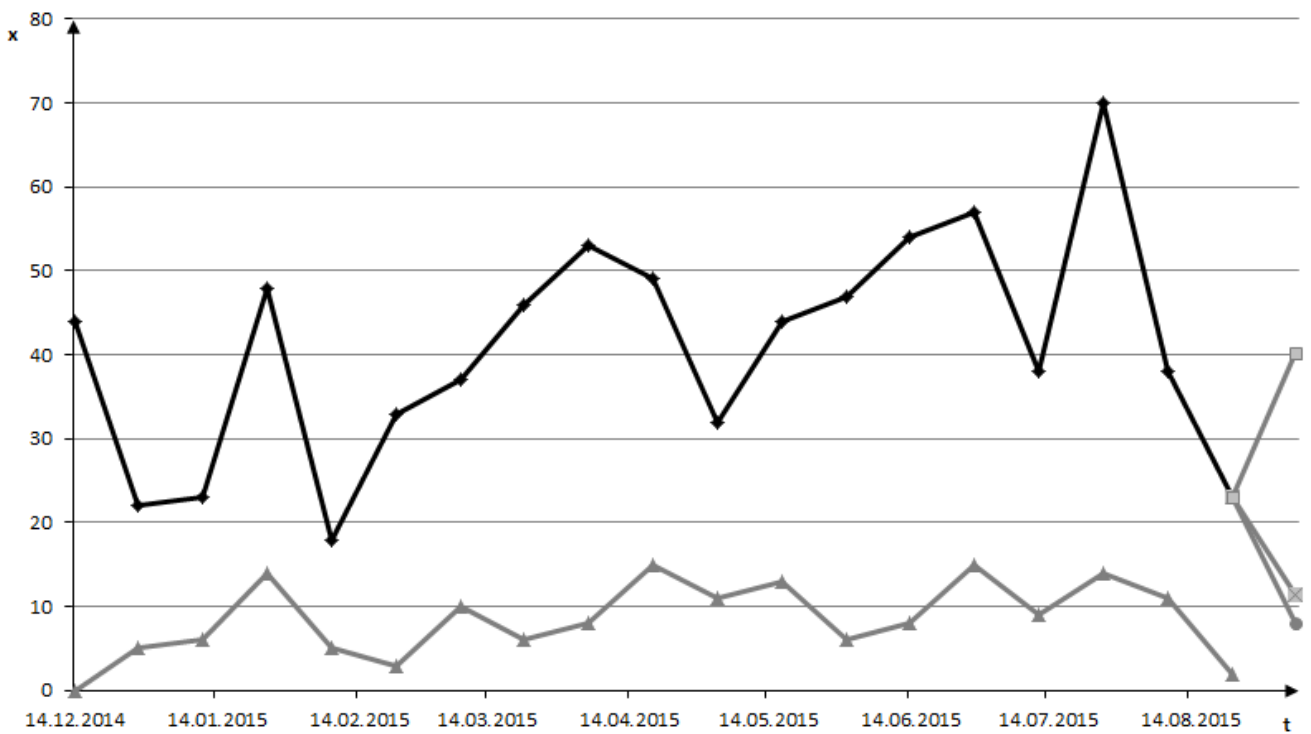


Рисунок 4.7 – Прогнозирование Improvement в проекте MongoDB с учетом влияния New Feature в проект (сохранения тенденции)

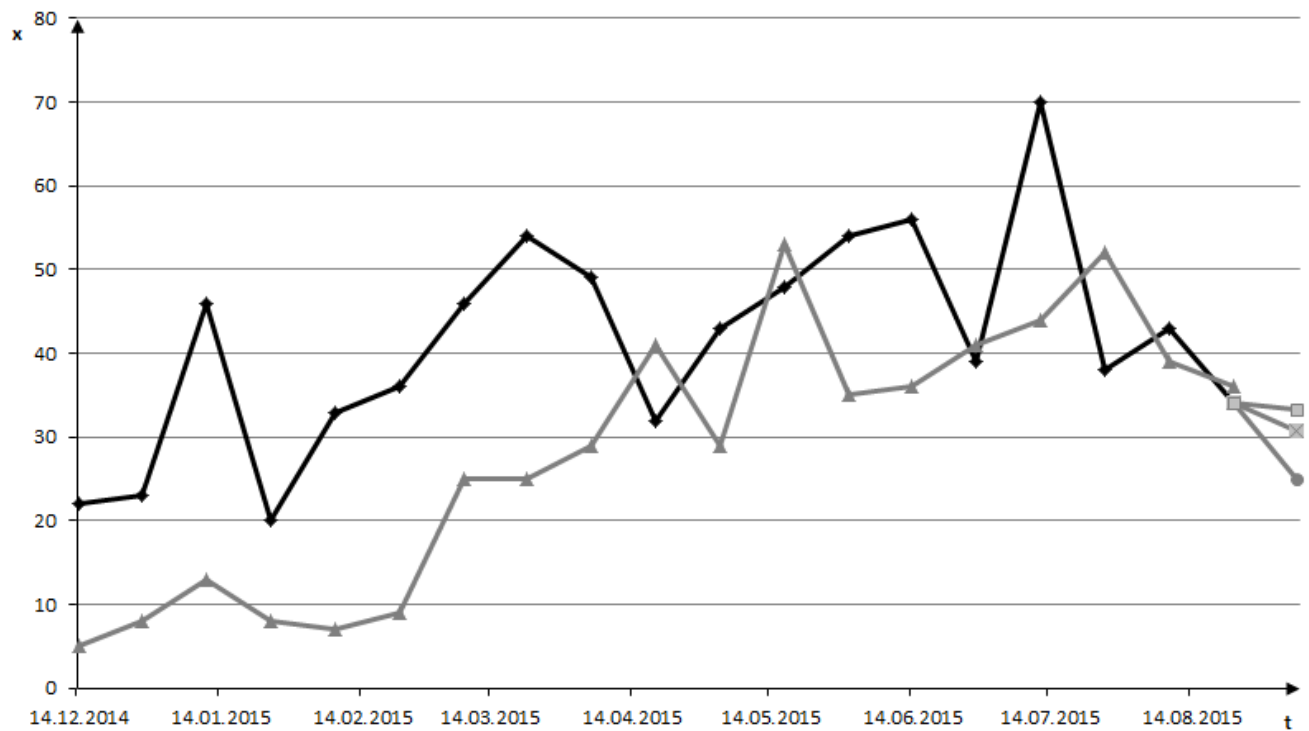


Рисунок 4.8 – Прогнозирование Improvement в проекте MongoDB с учетом влияния Task в проект (сохранения тенденции)

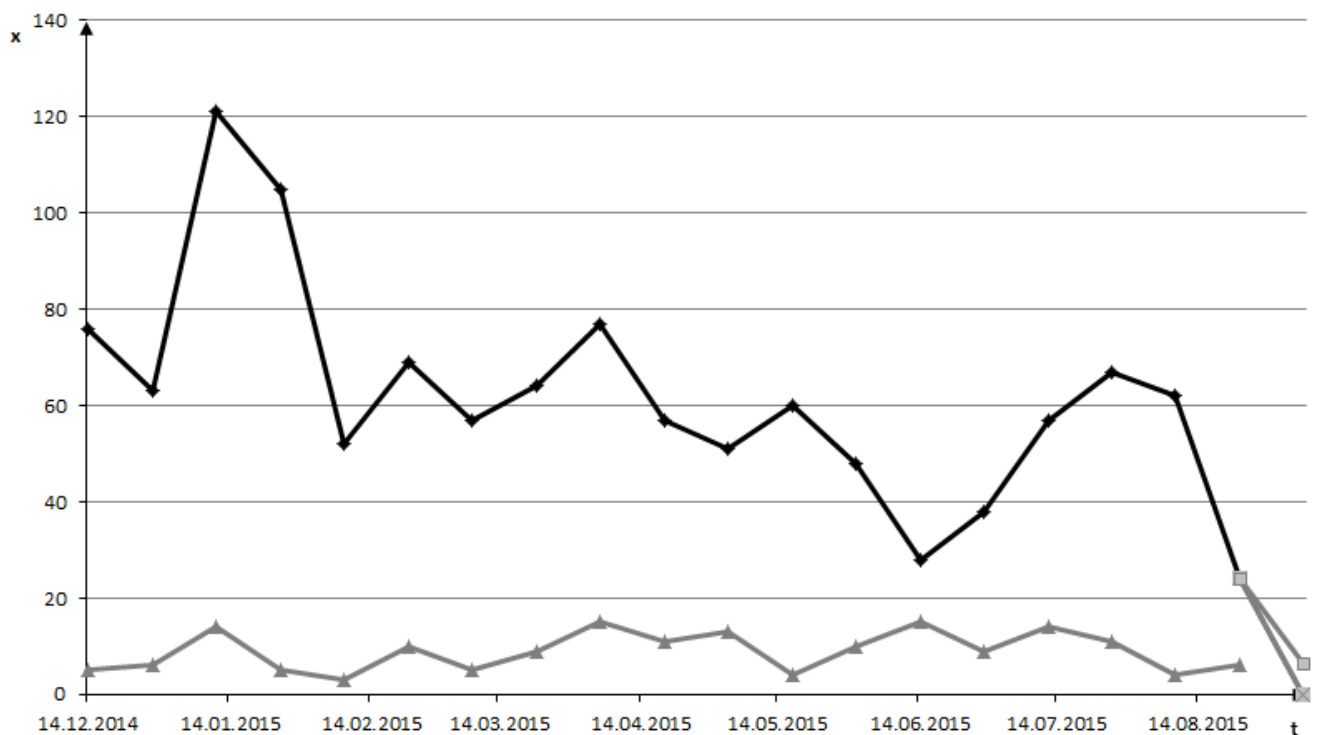


Рисунок 4.9 – Прогнозирование Bug в проекте MongoDB с учетом влияния Improvement в проект (сохранения тенденции)

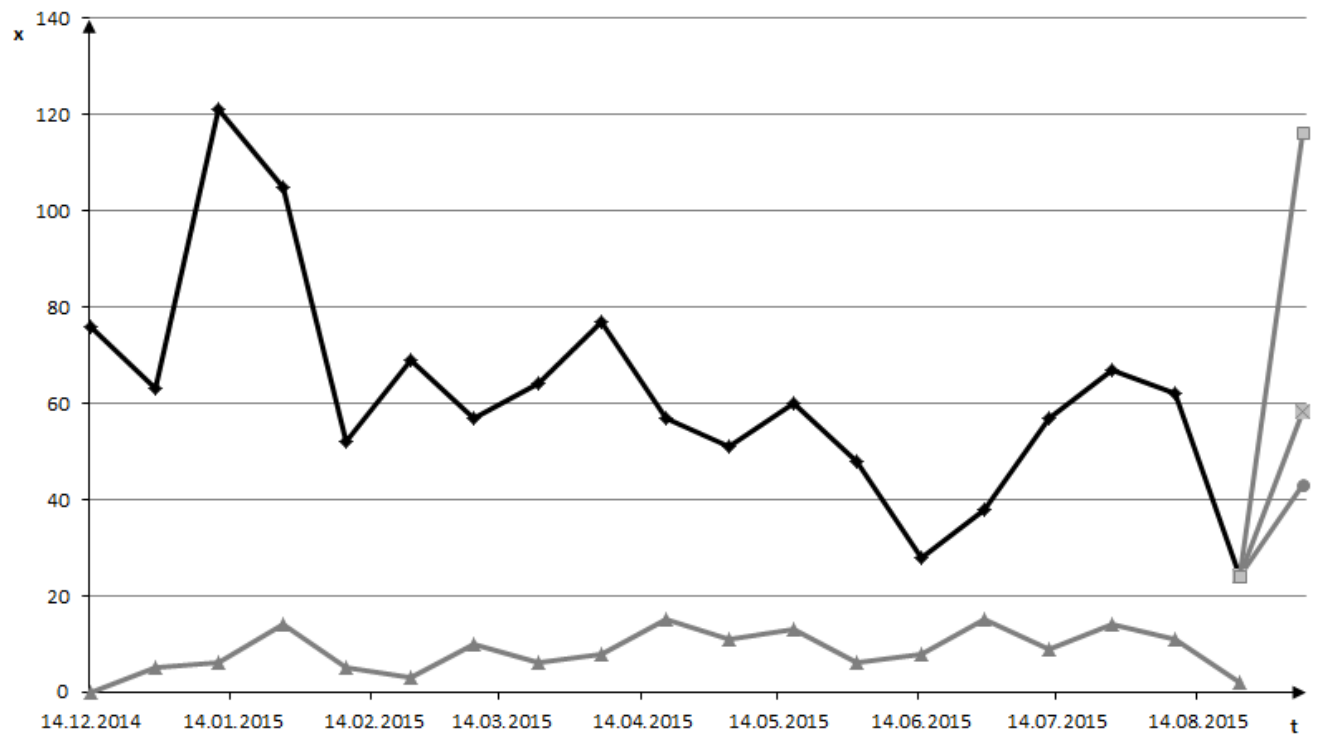


Рисунок 4.10 – Прогнозирование Bug в проекте MongoDB с учетом влияния New Feature в проект (гипотеза устойчивости тенденции)

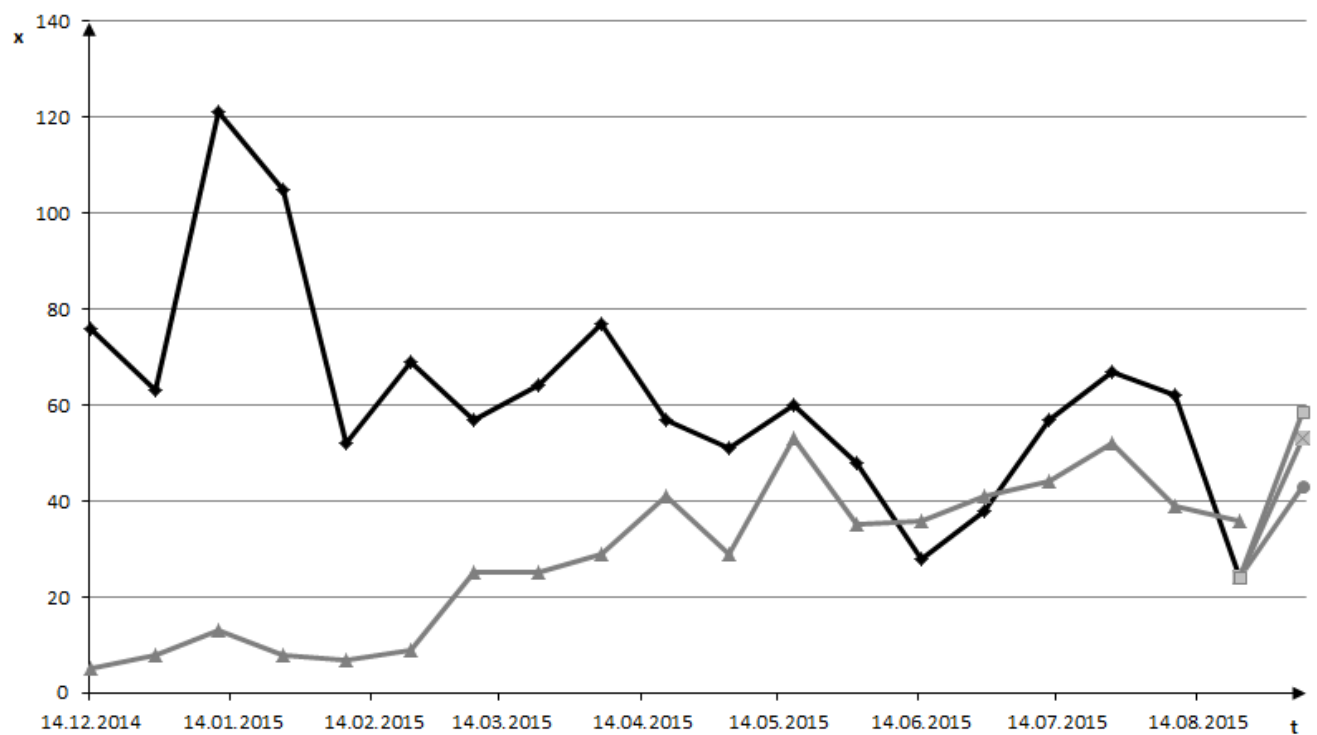


Рисунок 4.11 – Прогнозирование Bug в проекте MongoDB с учетом влияния Task в проект (гипотеза устойчивости тенденции)

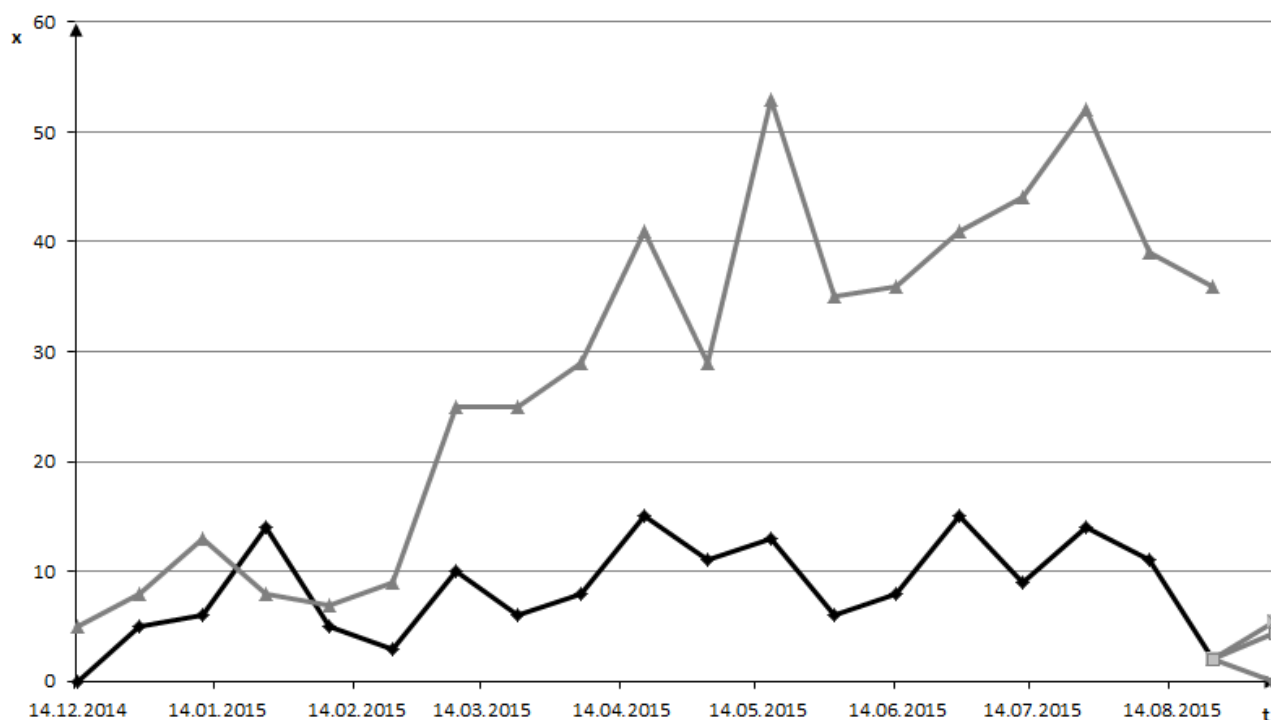


Рисунок 4.12 – Прогнозирование New Feature в проекте MongoDB с учетом влияния Task в проект (гипотеза на период)

Результаты оценки качества прогнозирования поступления документов в архив изделий с учетом влияния временных рядов предиктора представлены в таблицах 4.14 – MSE , в таблице 4.15 – $RMSE$. В Приложение G показаны результаты исследований пар ВР поступлений документов таблицы 4.2.

Результаты оценок качества прогноза метрик проектной деятельности – таблицы 4.16 (MSE) и 4.17 ($RMSE$).

Таблица 4.14 – Оценки качества прогноза критерием качества MSE

ВР поступления документов в архив (Код изделия)	ВР- предиктор (Код изделия)	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
14	54	0,00022	0,00018	0,0072	0,00081	0,00074	0,0069	0,0001	0,00061	0,0055
142	275	0,0007	0,0008	0,0008	0,0127	0,0123	0,0122	0,0344	0,0338	0,0336
14	220	0	0,0002	0,0009	0	0,0019	0,0093	0,0001	0,0001	0
54	220	0,0002	0,0002	0,0002	0	0,00001	0,00003	0	0,00001	0,00003
180	41	0,00002	0,00001	0,00001	0,00001	0,00002	0	0	0,00001	0,00001

Таблица 4.15 – Оценки качества прогноза критерием качества *RMSE*

ВР поступления документов в архив (Код изделия)	ВР- предиктор (Код изделия)	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
14	54	0	0,003	0,0847	0	0,0013	0,0829	0,0087	0,0057	0,0742
142	275	0,0265	0,0281	0,0286	0,1126	0,111	0,1105	0,1854	0,1839	0,1834
14	220	0	0,003	0,0019	0	0,0013	0,0019	0,0087	0,0075	0,0067
54	220	0,0139	0,014	0,0148	0	0,0001	0,0009	0	0,0001	0,0009
180	41	0,0054	0,0056	0,0059	0,0054	0,0052	0,0049	0,0054	0,0052	0,0049

Таблица 4.16 – Оценки качества прогноза критерием качества *MSE*

ВР метрик	ВР-предиктор	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
<i>Improvement</i>	<i>Feature</i>	0,0051	0,0051	0,1725	0,1961	0,1961	0,6189	0,059	0,059	0,3442
<i>Improvement</i>	<i>Task</i>	0,0051	0,0051	0,0003	0,1961	0,1961	0,1503	0,059	0,059	0,0352
<i>Bug</i>	<i>Improvement</i>	0,0017	0,0158	0,0019	0,0247	0,0053	0,0239	0,0055	0,0252	0,0059
<i>Bug</i>	<i>Feature</i>	0,0017	0,0158	0,0476	0,0247	0,0053	0,1735	0,0055	0,0252	0,0343
<i>Bug</i>	<i>Task</i>	0,0017	0,0158	0,0327	0,0247	0,0053	0,0003	0,0055	0,0252	0,0457
<i>Feature</i>	<i>Task</i>	0,2178	0,0148	0,0313	0,1111	0,4599	0,3882	0,0711	0,0061	0,0005

Таблица 4.17 – Оценки качества прогноза критерием качества *RMSE*

ВР метрик	ВР-предиктор	Гипотеза 1			Гипотеза 2			Гипотеза 3		
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
<i>Improvement</i>	<i>Feature</i>	0,0714	0,0714	0,4153	0,4429	0,4429	0,7867	0,2429	0,2429	0,5867
<i>Improvement</i>	<i>Task</i>	0,0714	0,0714	0,0163	0,4429	0,4429	0,3877	0,2429	0,2429	0,1877
<i>Bug</i>	<i>Improvement</i>	0,0413	0,1257	0,0439	0,157	0,0727	0,1545	0,0744	0,1587	0,0769
<i>Bug</i>	<i>Feature</i>	0,0413	0,1257	0,2182	0,157	0,0727	0,4166	0,0744	0,1587	0,1852
<i>Bug</i>	<i>Task</i>	0,0413	0,1257	0,1808	0,157	0,0727	0,0176	0,0744	0,1587	0,2138
<i>Feature</i>	<i>Task</i>	0,4667	0,1218	0,1769	0,3333	0,6782	0,6231	0,2667	0,0782	0,0231

4.4.1 «Измерение» метрик программных проектов с открытым кодом

Открытый проект «FreeNAS10». Взяты ВР «закрытых» «Bug» и «Feature» [10]. В таблицах 4.18, 4.19 представлены результаты анализа и прогнозирования метрик проектной деятельности.

Таблица 4.18 – Анализ данных проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
<i>Bug</i>	Падение	Стабильность	0,82	0,9736	Сильная
<i>New Feature</i>	Падение	Стабильность			

Таблица 4.19 – Результаты прогнозирования с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
<i>Bug</i>	<i>New Feature</i>	Стабильность	Рост средний	Падение сильное

В таблице 4.20 представлен прогноз появления «Bug». На рисунке 4.13 прогноз показан графически.

Таблица 4.20 – Значения метрики «Bug» по результатам прогноза

Гипотеза	Значение метрики «Bug»			Тенденция, интенсивность ($\hat{\tau}_{t+1}^Y$)
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	
Гипотеза сохранения тенденции	7	5	3	Стабильность
Гипотеза устойчивости тенденции	15	13	11	Рост средний
Гипотеза прогноза на заданный период	5	3	1	Падение сильное

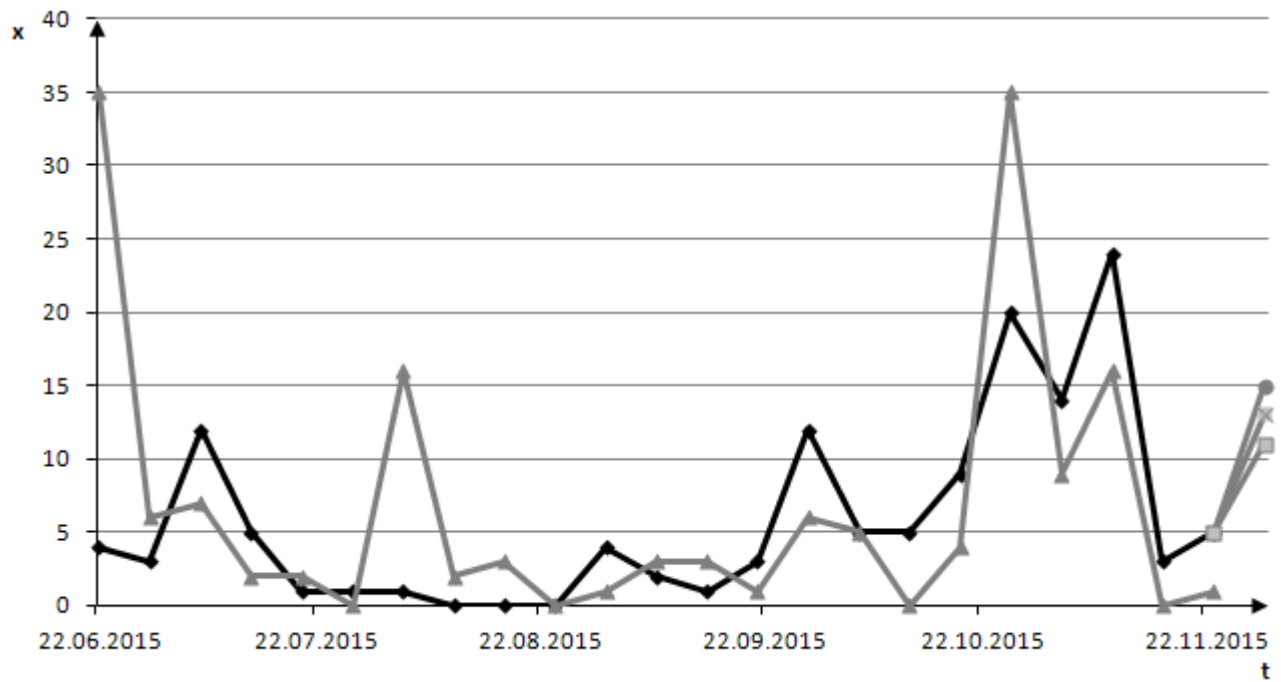


Рисунок 4.13 – Прогнозирование появления ошибок в проекте «FreeNAS10» с учетом влияния добавления новых функциональных возможностей в проект (гипотеза учета устойчивости тенденции)

Результаты оценки качества прогнозирования ошибок в разработке проекта с учетом влияния добавленных новых функций в проект представлены в таблице 4.21.

Таблица 4.21 – Оценка качества прогноза («Bug-New Feature»)

Гипотеза	MSE			RMSE		
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Гипотеза сохранения тенденции	0,7656	0,765	1,5625	0,875	0,8747	1,25
Гипотеза устойчивости тенденции	0,1914	0,2104	0,006	0,4375	0,4587	0,0774
Гипотеза прогноза на заданный период	0,0017	0,0018	0,1151	0,0417	0,042	0,3392

Открытый проект «FreeNAS9». Взяты закрытые ВР «Bug» и «Feature» [11]. В таблицах 4.22, 4.23 представлены результаты анализа и прогнозирования метрик проектной деятельности.

Таблица 4.22 – Анализ данных проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
<i>Bug</i>	Рост	Стабильность	0,75	0,9387	Сильная
<i>New Feature</i>	Рост	Стабильность			

Таблица 4.23 – Результаты прогнозирования с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
<i>Bug</i>	<i>New Feature</i>	Падение среднее	Рост сильный	Падение среднее

В таблице 4.24 представлен прогноз появления «Bug». Графически прогноз изображен на рисунке 4.14.

Таблица 4.24 – Значения метрики «Bug» по результатам прогноза

Гипотеза	Значение «Bug»			Тенденция, интенсивность ($\hat{\tau}_{t+1}^Y$)
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	
Гипотеза сохранения тенденции	0	0	0	Падение среднее
Гипотеза устойчивости тенденции	567	584	600	Рост сильный
Гипотеза прогноза на заданный период	128	145	161	Падение среднее

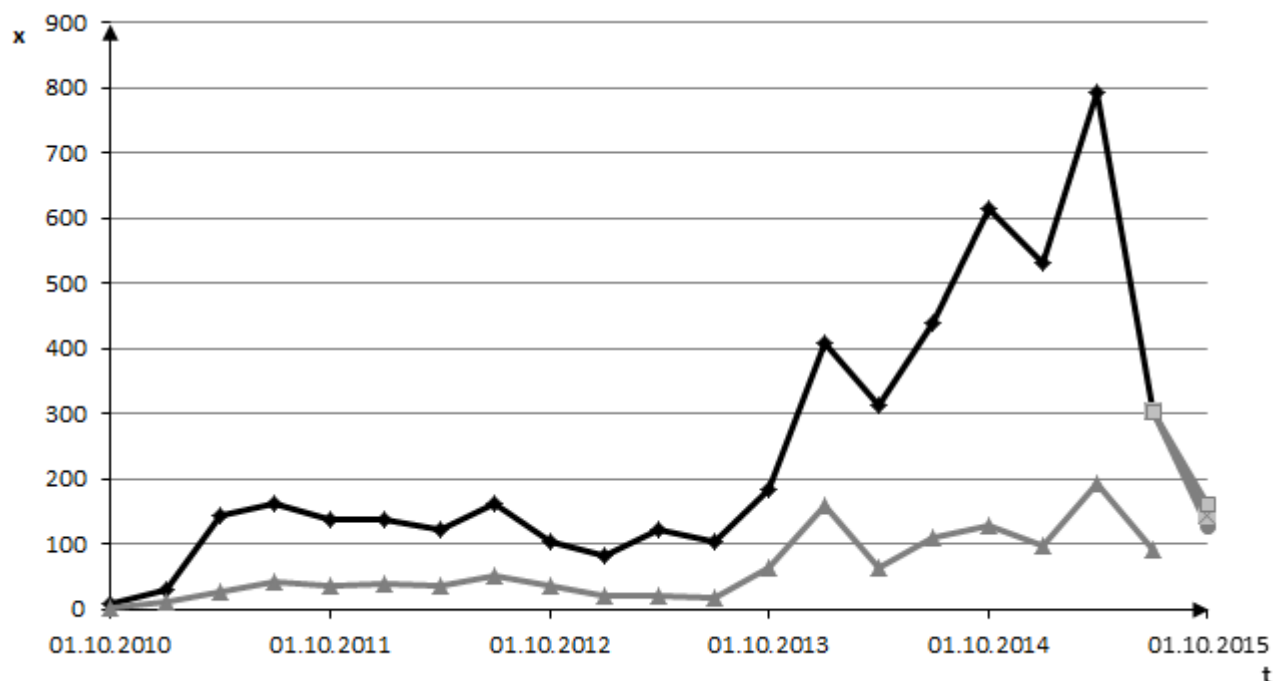


Рисунок 4.14 – Прогнозирование появления ошибок в проекте «FreeNAS9» с учетом влияния добавления новых функциональных возможностей в проект (гипотеза прогноза на заданный период)

Результаты оценки качества прогнозирования ошибок в разработке проекта с учетом влияния добавленных новых функций в проект представлены в таблице 4.25.

Таблица 4.25 – Оценка качества прогноза («Bug - New Feature»)

Гипотеза	MSE			RMSE		
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Гипотеза сохранения тенденции	0,1097	0,0937	0,1042	0,3312	0,3061	0,3227
Гипотеза устойчивости тенденции	0,1097	0,0937	0,1042	0,3312	0,3061	0,3227
Гипотеза прогноза на заданный период	0,0003	0,0001	0,0001	0,0164	0,0087	0,0079

Открытый проект «*Open Panel*». Взяты закрытые ВР «*Bug*» и «*Feature*» [30]. В таблицах 4.26 – 4.27 представлены результаты анализа и прогнозирования метрик проекта.

Таблица 4.26 – Анализ данных проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
<i>Bug</i>	Падение	Падение слабое	0,6	0,4464	Умеренная
<i>New Feature</i>	Падение	Падение стабильное			

Таблица 4.27 – Результаты прогнозирования с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
<i>Bug</i>	<i>New Feature</i>	Падение среднее	Падение среднее	Падение среднее

В таблице 4.28 представлен прогноз метрики «*Bug*». На рисунке 4.15 прогноз показан графически.

Таблица 4.28 – Значения «*Bug*» по результатам прогноза

Гипотеза	Значение метрики « <i>Bug</i> »			Тенденция, интенсивность ($\hat{t}_{t+1}^Y$)
	τ_{t+1}	τ_{t+1}^Y	$\hat{t}_{t+1}^Y$	
Гипотеза сохранения тенденции	1	0	0	Падение среднее
Гипотеза устойчивости тенденции	5	0	0	Падение среднее
Гипотеза прогноза на заданный период	1	0	0	Падение среднее

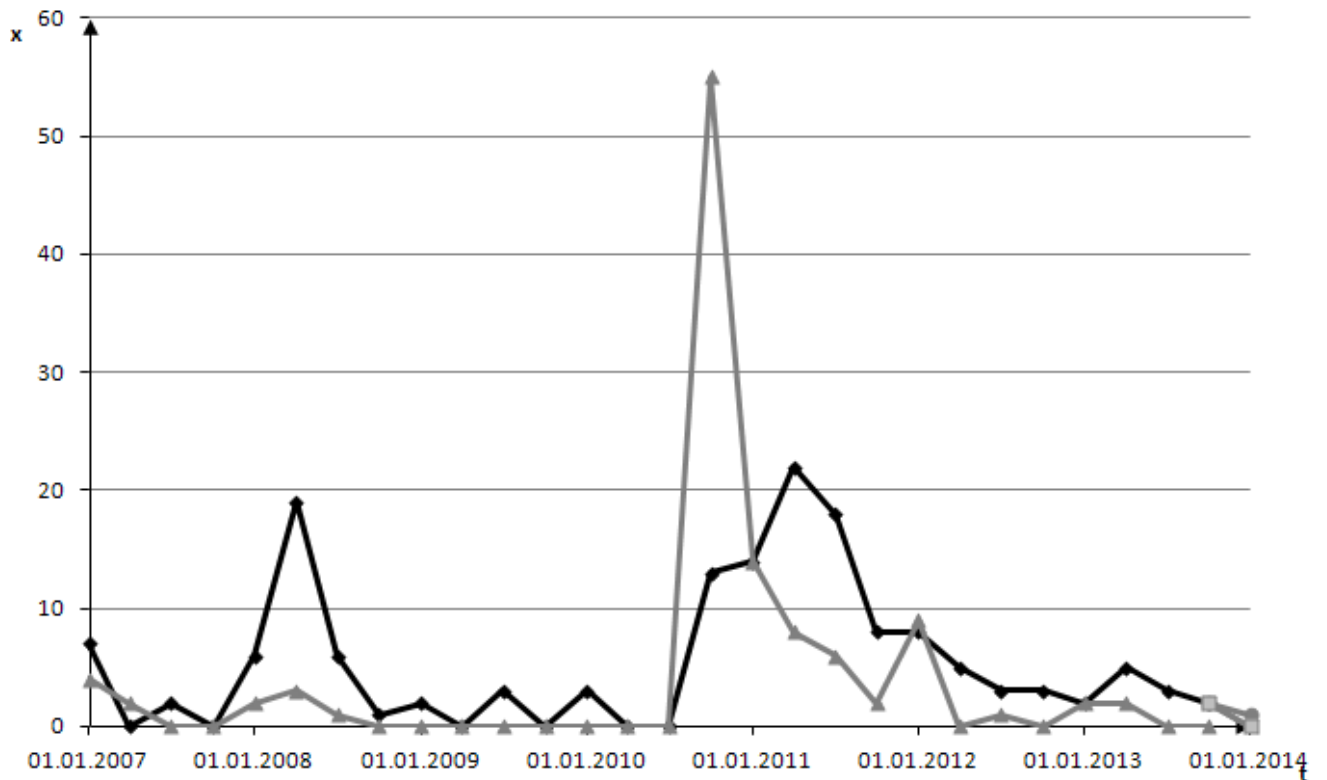


Рисунок 4.15 – Прогнозирование появления ошибок в проекте «*Open Panel*» с учетом влияния добавления новых функциональных возможностей в проект (гипотеза сохранения тенденции)

Результаты оценки качества прогнозирования ошибок в разработке проекта с учетом влияния добавленных новых функций в проект представлены в таблице 4.29.

Таблица 4.29 – Оценка качества прогноза («*Bug - New Feature*»)

Гипотеза	MSE			$RMSE$		
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Гипотеза сохранения тенденции	0,0083	0,0006	0,0004	0,0909	0,0236	0,0195
Гипотеза устойчивости тенденции	0,0186	0,0629	0,0609	0,1304	0,2509	0,2468
Гипотеза прогноза на заданный период	0,0186	0,0629	0,0609	0,1364	0,2509	0,2468

Открытый проект «PCBSD». Взяты закрытые ВР «*Bug*» и «*Feature*» [31]. В таблицах 4.30 – 4.31 представлены результаты анализа и прогнозирования метрик проекта.

Таблица 4.30 – Анализ данных проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
<i>Bug</i>	Рост	Рост слабый	0,6	0,6128	Средняя
<i>New Feature</i>	Рост	Стабильность			

Таблица 4.31 – Результаты прогнозирования с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
<i>Bug</i>	<i>New Feature</i>	Стабильность	Рост средний	Падение сильное

В таблице 4.32 представлен прогноз метрики «*Bug*». Графически прогноз представлен на рисунке 4.16.

Таблица 4.32 – Значения метрики «*Bug*» по результатам прогноза

Гипотеза	Значение метрики « <i>Bug</i> »			Тенденция, интенсивность ($\hat{\tau}_{t+1}^Y$)
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	
Гипотеза сохранения тенденции	32	32	31	Стабильность
Гипотеза устойчивости тенденции	37	37	36	Рост средний
Гипотеза прогноза на заданный период	20	20	19	Падение сильное

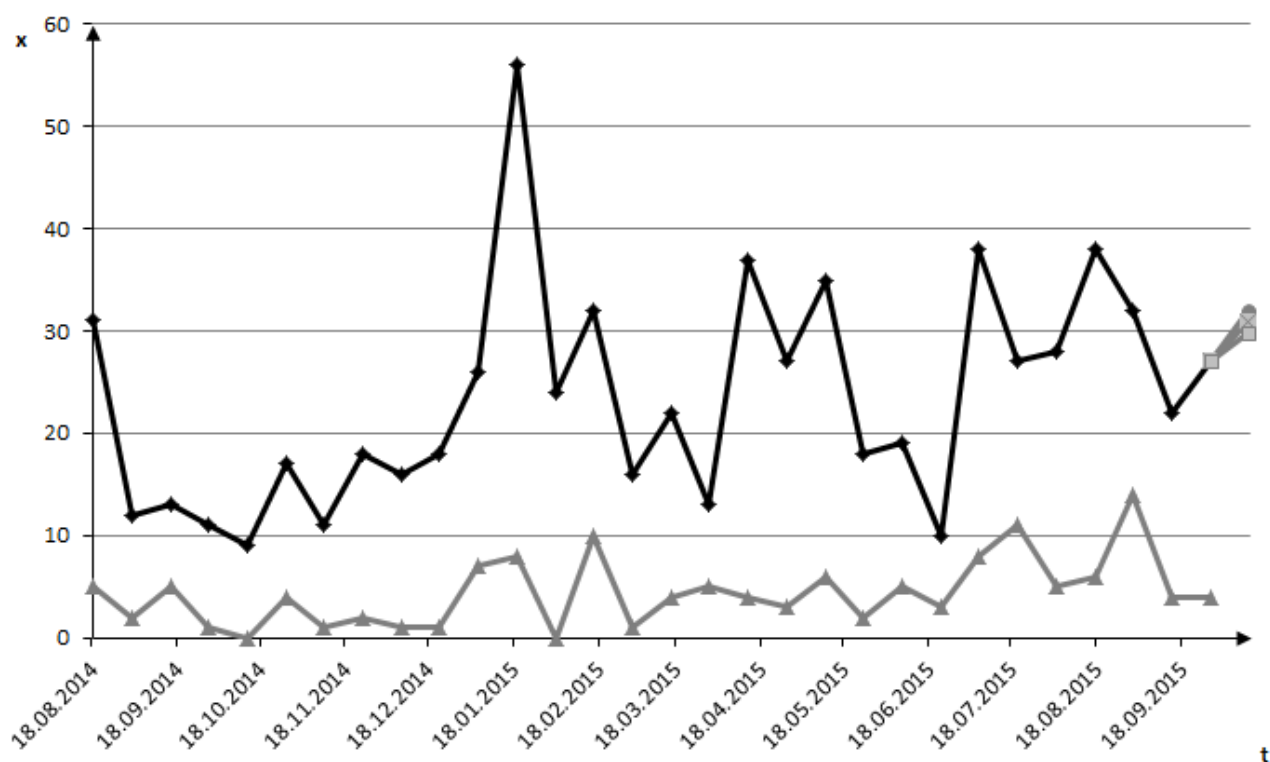


Рисунок 4.16 – Прогнозирование появления ошибок в проекте «PCBSD» с учетом влияния добавления новых функциональных возможностей в проект (гипотеза сохранения тенденции)

Результаты оценки качества прогнозирования ошибок в разработке проекта с учетом влияния добавленных новых функций в проект представлены в таблице 4.33.

Таблица 4.33 – Оценка качества прогноза («Bug- New Feature»)

Гипотеза	MSE			RMSE		
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Гипотеза сохранения тенденции	0,0319	0,0037	0,0015	0,1786	0,0611	0,0385
Гипотеза устойчивости тенденции	0,0319	0,0877	0,0982	0,1786	0,2961	0,3133
Гипотеза прогноза на заданный период	0,0013	0,0309	0,0291	0,0357	0,1758	0,1705

Проект по разработке автоматизированной системы «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов» по заказу ООО «Эверест Ресерч». Взяты ВР «*Bug*» и «*Feature*»(Приложение F). В таблицах 4.34 – 4.35 представлены результаты анализа и прогнозирования проектной деятельности.

Таблица 4.34 – Анализ данных проекта

Тип метки	Общая тенденция	Доминирующая тенденция	Мера сходства	Корреляция	Интерпретация корреляции
<i>Bug</i>	Рост	Рост слабый	0,59	0,4903	Умеренная
<i>New Feature</i>	Рост	Рост слабый			

Таблица 4.35 – Результаты прогнозирования с учетом влияния ВР-предиктора

ВР зависимый	ВР-предиктор	Гипотеза 1	Гипотеза 2	Гипотеза 3
<i>Bug</i>	<i>New Feature</i>	Падение сильное	Падение сильное	Падение среднее

В таблице 4.36 и на рисунке 4.17 представлен прогноз метрики «*Bug*».

Таблица 4.36 – Значения метрики «*Bug*» по результатам прогноза

Гипотеза	Значение метрики « <i>Bug</i> »			Тенденция, интенсивность ($\hat{\tau}_{t+1}^Y$)
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	
Гипотеза сохранения тенденции	7	6	5	Падение сильное
Гипотеза устойчивости тенденции	7	6	5	Падение сильное
Гипотеза прогноза на заданный период	6	5	4	Падение среднее

Результаты оценки качества прогнозирования ошибок в разработке проекта с учетом влияния добавленных новых функций в проект представлены в таблице 4.37.

Гипотеза	MSE			$RMSE$		
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$
Гипотеза сохранения тенденции	0,1111	0,0275	0,0275	0,3333	0,1659	0,1659
Гипотеза устойчивости тенденции	0,25	0,4455	0,4455	0,5	0,6674	0,6674
Гипотеза прогноза на заданный период	0,4444	0,6957	0,6957	0,6677	0,8341	0,8347

Согласно предложенному алгоритму прогнозирования на первой фазе при прогнозировании по гипотезе сохранения тенденции ВР1 У (14 код изделия) (рис. 4.2) была получена прогнозная НЭТ «Стабильность» и прогнозные значения

оценки качества ($MSE=0,00022$) , затем, на второй фазе, применяя алгоритм 1, учитывалась основная доминирующая НТ BP1 Y «Стабильность» (табл. 4.1) и доминирующая НТ BP2 Z (54 код изделия) «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE=0,0072$) (см. табл. 4.14). При корректировке прогнозирования по гипотезе устойчивости ошибка MSE изменилась с 0,00081 до 0,0069. При использовании гипотезы прогнозирования на период MSE изменилась в большую сторону с 0,0001 до 0,0055.

В эксперименте проектов с кодом изделия 142 и 275 на первой фазе предложенного алгоритма при прогнозировании по гипотезе сохранения тенденции BP1 Y (142 код изделия) (рис. 4.3) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0,0007$) , затем, на второй фазе применяя алгоритм 1 учитывалась основная доминирующая НТ BP1 Y «Стабильность» (табл. 4.1), общей тенденции «Рост» и доминирующая НТ BP2 Z (275 код изделия) «Стабильность» при общей тенденции «Падение», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE=0,0008$) (см. табл. 4.14). При адаптации прогнозирования по гипотезе устойчивости ошибка MSE уменьшилась с 0,0127 до 0,0122. При использовании гипотезы прогнозирования на период MSE изменилась в меньшую сторону с 0,0344 до 0,0336.

Эксперимент с временными рядами поступления отчетов по изделиям 14 и 220 показал, что на первой фазе при прогнозировании по гипотезе сохранения тенденции BP1 Y (14 код изделия) (рис. 4.4) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0$), затем, на второй фазе применяя алгоритм 1 была взята полученная ранее основная доминирующая НТ BP1 Y «Стабильность» (табл. 4.1) и доминирующая НТ BP2 Z (220 код изделия) «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE=0,0009$) (см. табл.

4.14). При корректировке прогноза по гипотезе устойчивости ошибка MSE увеличилась с 0 до 0,0093. При использовании гипотезы прогнозирования на период MSE изменилась в меньшую сторону с 0,0001 до 0.

При прогнозировании по гипотезе сохранения тенденции в эксперименте с изделиями 54 и 220 на первой фазе предложенного алгоритма BP1 Y (54 код изделия) (рис. 4.5) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0,0002$), затем, на второй фазе, применяя алгоритм 1, была учтена полученная ранее основная доминирующая НТ BP1 Y «Стабильность» (табл. 4.1) и доминирующая НТ BP2 Z (220 код изделия) «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых использовались для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,0002$) (см. табл. 4.14). При адаптации прогноза по гипотезе устойчивости ошибка MSE увеличилась с 0 до 0,00003. При использовании гипотезы прогнозирования на период MSE так же изменилась в большую сторону с 0 до 0,00003.

Временные ряды поступления документации в архив проектов с кодами изделия 180 и 41 при прогнозировании по гипотезе сохранения тенденции согласно предлагаемому алгоритму в результате экспериментов показали, что на первой фазе BP1 Y (180 код изделия) (рис. 4.6) прогнозная НЭТ имеет значение «Стабильность» и прогнозные значения оценки качества ($MSE=0,00002$), затем, на второй фазе, применяя алгоритм 1, учитывалась основную доминирующую НТ BP1 Y «Стабильность» (табл. 4.1) и доминирующая НТ BP2 Z (41 код изделия) «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,00001$) (см. табл. 4.14). При коррекции прогнозирования по гипотезе устойчивости ошибка MSE уменьшилась с 0,00001 до 0. При использовании гипотезы прогнозирования на период MSE изменилась в большую сторону с 0 до 0,00001.

В результате экспериментов по гипотезе сохранения тенденции при прогнозировании метрик проекта *MongoDB* согласно предлагаемому алгоритму

на первой фазе BP1 Y («*Improvement*») (рис. 4.7) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0,0051$), затем, на второй фазе применяя алгоритм 1 была учтена основная доминирующая НТ BP1 Y «Рост слабый» (табл. 4.2) и доминирующая НТ BP2 Z («*New Feature*») «Рост средний» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,1725$) (см. табл. 4.16). При корректировке прогноза по гипотезе устойчивости ошибка MSE увеличилась с 0,1961 до 0,6189. При использовании гипотезы прогнозирования на период MSE так же увеличилась с 0,059 до 0,3442.

В эксперименте с метриками проекта «*Improvement*» и «*Task*» на первой фазе предложенного алгоритма при прогнозировании по гипотезе сохранения тенденции BP1 Y («*Improvement*») (рис. 4.8) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0,0051$), затем, на второй фазе, применяя алгоритм 1, учитывалась основная доминирующая НТ BP1 Y «Рост слабый» (табл. 4.2), общей тенденции «Рост» и доминирующая НТ BP2 Z («*Task*») «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE=0,0003$) (см. табл. 4.16). При адаптации прогнозирования по гипотезе устойчивости ошибка MSE уменьшилась с 0,1961 до 0,1503. При использовании гипотезы прогнозирования на период MSE изменилось в меньшую сторону с 0,059 до 0,0352.

Эксперимент с временными рядами метрик проектной деятельности «*Bug*» и «*Improvement*» показал, что на первой фазе при прогнозировании по гипотезе сохранения тенденции BP1 Y («*Bug*») (рис. 4.9) была получена прогнозная НЭТ «Стабильность» и прогнозные значения оценки качества ($MSE=0,0017$), затем, на второй фазе, применяя алгоритм 1, была взята полученная ранее основная доминирующая НТ BP1 Y «Рост слабый» (табл. 4.2) и доминирующая НТ BP2 Z («*Improvement*») «Рост слабый» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции

прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,0019$) (см. табл. 4.16). При адаптации прогноза по гипотезе устойчивости ошибка MSE уменьшилась с 0,0247 до 0,0239. При использовании гипотезы прогнозирования на период MSE увеличилась с 0,0055 до 0,0059.

При прогнозировании по гипотезе сохранения тенденции в эксперименте с метриками проекта «Bug» и «New Feature» на первой фазе предложенного алгоритма BP1 Y («Bug») (рис. 4.10) была получена прогнозная НЭТ «Рост средний» и прогнозные значения оценки качества ($MSE=0,0017$), затем, на второй фазе, применяя алгоритм 1, была учтена полученная ранее основная доминирующая НТ BP1 Y «Рост слабый» (табл. 4.2) и доминирующая НТ BP2 Z («New Feature») «Рост средний» при общей тенденции «Рост», дефаззифицированные значения которых использовались для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,0017$) (см. табл. 4.16). При корректировке прогноза по гипотезе устойчивости ошибка MSE увеличилась с 0,0247 до 0,1735. При использовании гипотезы прогнозирования на период MSE так же изменилась в большую сторону с 0,0055 до 0,0343.

Временные ряды метрик проектной деятельности «Bug» и «Task» при прогнозировании по гипотезе сохранения тенденции согласно предлагаемому алгоритму в результате экспериментов показали, что на первой фазе BP1 Y («Bug») (рис. 4.11) прогнозная НЭТ имеет значение «Падение среднее» и прогнозные значения оценки качества ($MSE=0,0017$), затем, на второй фазе, применяя алгоритм 1, учитывалась основная доминирующая НТ BP1 Y «Рост слабый» (табл. 4.2) и доминирующая НТ BP2 Z («Task») «Стабильность» при общей тенденции «Рост», дефаззифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,0327$) (см. табл. 4.16). При адаптации прогноза по гипотезе устойчивости ошибка MSE уменьшилась с 0,0247 до 0,0003. При использовании гипотезы прогнозирования на период MSE изменилась в большую сторону с 0,0055 до 0,0457.

Согласно предложенному алгоритму прогнозирования на первой фазе при прогнозировании по гипотезе сохранения тенденции BP1 Y («*New Feature*») (рис. 4.12) была получена прогнозная НЭТ «Рост сильный» и прогнозные значения оценки качества ($MSE=0,2178$), затем, на второй фазе, применяя алгоритм 1, учитывалась основная доминирующая НТ BP1 Y «Рост средний» (табл. 4.2) и доминирующая НТ BP2 Z («*Task*») «Стабильность» при общей тенденции «Рост», дефазифицированные значения которых были использованы для коррекции прогнозных значений BP1 Y полученного на первой фазе ($MSE= 0,0313$) (см. табл. 4.16). При корректировке прогноза по гипотезе устойчивости ошибка MSE увеличилась с 0,1111 до 0,3882. При использовании гипотезы прогнозирования на период MSE изменилась в меньшую сторону с 0,0711 до 0,0005.

Полученные результаты всех экспериментов показывают, что предложенный подход, реализующий модифицированный метод прогнозирования НЭТ может быть использован для краткосрочного прогнозирования ВР в ситуации, когда имеется экспертное предположение о существовании ВР предиктора.

В эксперименте с временными рядами поступления отчетов по код изделия 142 и 275 было показано, что предложенный подход релевантен гипотезам устойчивости тенденции и при «прогнозе на заданный период», поскольку в результате прогнозирования получили тенденцию «Рост слабый», на которую повлияла общая тенденция предиктора «Падение». При прогнозировании с помощью гипотезы сохранения тенденции метод корректировки менее актуален.

Эксперимент с временными рядами поступления отчетов по код изделия 14 и 220 продемонстрировал использование рекомендованного подхода в гипотезе прогноза на заданный период. При прогнозировании с помощью гипотезы сохранения и устойчивости тенденции метод корректировки менее значим.

Результаты в ходе эксперимента над изделиями с кодами 54 и 220 не подтвердили важность использования предлагаемого подхода для краткосрочного прогнозирования ВР с учетом влияния ВР-предиктора. Но актуальность

использования гипотез без корректировки доказывается низким значением полученных ошибок.

Для краткосрочного получения прогноза временных рядов поступления отчетов документации проектов с кодами изделий 180 и 41 с экспертным заключением о наличии ВР-предиктора актуально использование рекомендованного подхода в гипотезе «прогноза на заданный период». При прогнозировании с помощью гипотезы сохранения и устойчивости тенденции метод корректировки дает незначительное улучшение.

В эксперименте с временными рядами метрик «*Improvement*» и «*New Feature*» проектной документации было показано, что предложенный подход наиболее подходит в гипотезах сохранения тенденции и прогнозе на заданный период, поскольку *MSE* составила не более 1-2% от среднего значения метрик прогноза. При прогнозировании с помощью гипотезы сохранения тенденции метод корректировки менее актуален.

Эксперимент с временными рядами метрик «*Improvement*» и «*Task*» продемонстрировал использования рекомендованного подхода во всех трех гипотезах, так как *MSE* составила до 2% от среднего значения метрик прогноза.

Результаты в ходе экспериментов над метриками «*Bug*» / «*Improvement*» и «*Bug*» / «*Task*» также указали на возможность использования предлагаемого подхода для краткосрочного прогнозирования ВР с учетом влияния ВР-предиктора во всех трех гипотезах.

Исследование метрик «*Bug*» и «*New Feature*» указали на не значимое использование предлагаемого подхода при корректировке гипотезы устойчивости тенденции. В остальных случаях результаты исследования показали, что *MSE* составила не более 2% от среднего значения метрик, что говорит о значимости предложенного подхода прогнозирования.

Для краткосрочного получения прогноза временных рядов метрик «*New Feature*» и «*Task*» проектной деятельности с экспертным заключением о наличии ВР-предиктора актуально использование рекомендованного подхода в гипотезе

сохранения тенденции и «прогноза на заданный период». При прогнозировании с помощью гипотезы устойчивости тенденции метод корректировки не так показателен.

Согласно проведенным экспериментам открытых проектов по метрикам «Bug» и «New Feature», в проекте «FreeNAS10» была выявлена сильная корреляция и сходство между метриками, что свидетельствует о влиянии на прогноз ВР-предиктора. Так как основная доминирующая НТ ВР1 Y «Стабильность» (табл. 4.18) и доминирующая НТ ВР2 Z («New Feature») «Стабильность» при общей тенденции «Падение», было получено меньшее значение оценки качества MSE – не более 2% от среднего значения метрик при корректировке прогноза гипотезы устойчивости тенденции (табл. 4.21), что говорит о значимости предложенного подхода прогнозирования. В остальных случаях результаты исследования показали, что MSE составила более 2 % от среднего значения метрик.

В эксперименте открытого проекта «FreeNAS9» между метриками «Bug» и «New Feature» была выявлена сильная корреляция и сходство, что свидетельствует о влиянии на прогноз ВР-предиктора. Поскольку основная доминирующая НТ ВР1 Y «Стабильность» (табл. 4.22) и доминирующая НТ ВР2 Z («New Feature») «Стабильность» при общей тенденции «Рост», было получено меньшее значение оценки качества MSE – не более 2% от среднего значения метрик при корректировке прогноза гипотезы на заданный период (табл. 4.25), что говорит о значимости предложенного подхода прогнозирования.

Между метриками «Bug» и «New Feature» в эксперименте открытого проекта «OpenPanel» данные по корреляции и мере сходства были умеренными, что говорит о наличии влияния на прогноз ВР-предиктора. Основная доминирующая НТ ВР1 Y «Падение слабое» (табл. 4.26) и доминирующая НТ ВР2 Z («New Feature») «Падение стабильное» при общей тенденции «Падение», меньшее значение оценки качества MSE было получено не более 2% от среднего значения метрик при адаптации прогноза гипотезы сохранения тенденции (табл. 4.29), что говорит о значимости предложенного подхода прогнозирования.

В исследовании метрик «*Bug*» и «*New Feature*» открытого проекта «*PCBSD*» была выявлена достаточная корреляция и мера сходства между ними, чтобы утверждать наличие влияния на прогноз ВР-предиктора. Так как доминирующая НТ ВР1 *Y* «Рост слабый» (табл. 4.30) и доминирующая НТ ВР2 *Z* («*New Feature*») «Стабильность» при общей тенденции «Рост», меньшее значение оценки качества *MSE* было получено не более 2% от среднего значения метрик при адаптации прогноза гипотезы сохранения тенденции (табл. 4.29), что говорит о значимости предложенного подхода прогнозирования.

В результате анализа метрик «*Bug*» и «*New Feature*» открытого проекта по разработке автоматизированной системы «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов» по заказу ООО «Эверест Ресерч» показатели корреляции и меры сходства были получены умеренными, что свидетельствует о наличии влияния на прогноз ВР-предиктора. Поскольку доминирующая НТ ВР1 *Y* «Рост слабый» (табл. 4.34) и доминирующая НТ ВР2 *Z* («*New Feature*») «Рост слабый» при общей тенденции «Рост», было получено меньшее значение оценки качества *MSE* – не более 2% от среднего значения метрик при адаптации прогноза гипотезы сохранения тенденции (табл. 4.37), что говорит о значимости предложенного подхода прогнозирования.

Результаты проведенных экспериментов сравнивались с результатами метода *ARIMA*, производящим прогноз при одинаковых условиях с предложенной методикой: равном количестве точек временного ряда, участвующих в обучении, длине прогнозного участка. И для метода *ARIMA* и для предлагаемой методики выбираются такие параметры, которые обеспечивают наивысшее качество прогнозирования.

Итоги сравнения представлены в таблице 4.38-4.40.

Таблица 4.38 – Сравнение оценок прогноза по поступлениям документации с *ARIMA*

ВР поступления документов в архив (Код изделия)	ВР-предиктор (Код изделия)	Гипотеза 1			Гипотеза 2			Гипотеза 3			ARIMA	Среднее значение метрик
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$		
14	54	0,00022	0,00018	0,0072	0,00081	0,00074	0,0069	0,0001	0,00061	0,0055	0,0094	1,67
142	275	0,0007	0,0008	0,0008	0,0127	0,0123	0,0122	0,0344	0,0338	0,0336	0,0269	10,85
14	220	0	0,0002	0,0009	0	0,0019	0,0093	0,0001	0,0001	0	0,1342	1,67
54	220	0,0002	0,0002	0,0002	0	0,00001	0,00003	0	0,00001	0,00003	0,9232	2,48
180	41	0,00002	0,00001	0,00001	0,00001	0,00002	0	0	0,00001	0,00001	0,0685	5,11

Таблица 4.39 – Сравнение оценок прогноза метрик проекта «*MongoDB*» с *ARIMA*

ВР метрик	ВР-предиктор	Гипотеза 1			Гипотеза 2			Гипотеза 3			ARIMA	Среднее значение метрик
		τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$		
<i>Improvement</i>	<i>New Feature</i>	0,0051	0,0051	0,1725	0,1961	0,1961	0,6189	0,059	0,059	0,3442	0,4623	41,37
<i>Improvement</i>	<i>Task</i>	0,0051	0,0051	0,0003	0,1961	0,1961	0,1503	0,059	0,059	0,0352	0,4623	41,37
<i>Bug</i>	<i>Improvement</i>	0,0017	0,0158	0,0019	0,0247	0,0053	0,0239	0,0055	0,0252	0,0059	0,0059	61,89
<i>Bug</i>	<i>New Feature</i>	0,0017	0,0158	0,0476	0,0247	0,0053	0,1735	0,0055	0,0252	0,0343	0,0059	61,89
<i>Bug</i>	<i>Task</i>	0,0017	0,0158	0,0327	0,0247	0,0053	0,0003	0,0055	0,0252	0,0457	0,0059	61,89
<i>New Feature</i>	<i>Task</i>	0,2178	0,0148	0,0313	0,1111	0,4599	0,3882	0,0711	0,0061	0,0005	0,56	8,89

Таблица 4.40 – Сравнение оценок прогноза метрик «Bug»/«Feature» открытых проектов с ARIMA

Проекты	Гипотеза 1			Гипотеза 2			Гипотеза 3			ARIMA	Среднее значение метрик
	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$		
<i>FreeNAS10</i>	0,7656	0,765	0,5625	0,1914	0,2104	0,006	0,0017	0,0018	0,1151	0,4074	5,82
<i>FreeNAS9</i>	0,1097	0,0937	0,1042	0,1097	0,0937	0,1042	0,0003	0,0001	0,0001	12,255	245,1
<i>OpenPanel</i>	0,0083	0,0006	0,0004	0,0186	0,0629	0,0609	0,0186	0,0629	0,0609	0,4986	5,54
<i>PCBSD</i>	0,0319	0,0037	0,0015	0,0319	0,0877	0,0982	0,0013	0,0029	0,0029	1,1715	23,43
ПС «Эверест Ресерч»	0,1111	0,0275	0,0275	0,25	0,4455	0,4455	0,4444	0,6957	0,6957	0,5397	2,57

4.6 Выводы

В четвертой главе описаны результаты экспериментов, проведенные на временных рядах данных о поступлении отчетов по изделиям в архив отдела технической документации ФНПЦ АО «НПО «Марс», ВР метрик программной части Core Server открытого проекта «*MongoDB*», представлены результаты исследований метрик открытых проектов «*FreeNAS10*», «*FreeNAS9*», «*OpenPanel*», «*PCBSD*», а так же метрик открытого проекта по разработке автоматизированной системы «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов» по заказу ООО «Эверест Ресерч».

Проведенные вычислительные эксперименты по исследованию эффективности предложенных методов и алгоритмов показали, что в половине случаев предложенный подход показывает улучшение качества прогнозирования по сравнению с моделями класса *ARIMA* по критерию *MSE*, который составляет не более 1-2% от среднего значения метрик прогноза.

ЗАКЛЮЧЕНИЕ

В ходе диссертационного исследования были получены следующие результаты:

1. Главным итогом диссертационной работы является создание моделей, методов, алгоритмов анализа и средств управления проектной деятельностью на основе метрических характеристик.
2. Выполнен сравнительный анализ современных методов и средств построения различных видов систем контроля версий, средств управления проектами, а также научный обзор методов анализа временных рядов.
3. Разработана формализованная модель состояния проектной деятельности на основе метрик проектов, загруженных в систему контроля версий проектного репозитория.
4. Разработано средство анализа состояния метрик сущностей СКВ проектов программного обеспечения как инструмент управления проектными работами организации.
5. Разработаны модели и алгоритмы прогнозирования нечетких временных рядов метрических характеристик проекта.
6. Разработаны модели и алгоритмы прогнозирования временных рядов метрик проекта на основании мер сходства.
7. Разработан комплекс программ моделирования как подсистема автоматизации СКВ.

Таким образом, в диссертации решена актуальная научно-техническая задача, имеющая важное значение для развития и применения в системах автоматизированного проектирования средств математического моделирования, численных методов, а именно: построение эффективного средства управления совокупностью проектов крупной проектной организации на основе анализа и моделирования состояния метрик проектов программно-аппаратных комплексов.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ASL-шкала	(Absolute & Comparative Linguistic) – специальная лингвистическая шкала
ВР	Временной ряд
ГИП	Главный инженер проекта
ИПХ	Информационно-поисковая характеристика
ЛУ	Лист утверждения
НВР	Нечеткий временной ряд
НТ	Нечеткая тенденция
НЭТ	Нечеткая элементарная тенденция
ОУП	Офис управления проектами
РГП	Рабочая группа проекта
РН	Руководитель направления
САПР	Система автоматизированного проектирования
СКВ	Система контроля версий
ТЗ	Техническое задание
ФЗ	Функциональный Заказчик
ЭИР	Электронный информационный ресурс

СПИСОК ЛИТЕРАТУРЫ

1. Anderson T.W. Statistical Analysis of Time Series. New York: John Wiley and Sons, 1971.
2. Armstrong J.S., Collopy F. Error measures for generalizing about forecasting methods: Empirical comparisons // International Journal of Forecasting. 1992. Vol. 8(1). pp. 69-80.
3. Bardossy A. Note on fuzzy regression // Fuzzy Sets and Systems. 1990. Vol. 7. pp. 65-75.
4. Batyrshin I., Herrera R., Sheremetov L., Suarez R. Moving approximations in time series data mining // Proc. of Int. Conference on Fuzzy Sets and Soft Computing in Economics and Finance, vol. I. June 2004. pp. 62-72.
5. Batyrshin I., Herrera-Avelar R., Sheremetov L., Panova A. Association networks in time series data mining // NAFIPS 2005. Soft Computing for Real World Applications. June 2005. pp. 754-759.
6. Batyrshin I., Sheremetov L. Percepti on Based Time Series Data Mining for Decision Making // IFSA'07 Theoretical Advances and Applications of Fuzzy Logic. pp. 209-219.
7. Chen S.M. Forecasting enrollments based on high-order fuzzy time series // Cybernetics and Systems: An International Journal. 2002. Vol. 33. pp. 1–16.
8. Continuous Integration [Электронный ресурс] // Martin Fowler: [сайт]. URL: <http://www.martinfowler.com/articles/continuousIntegration.html#KeepTheBuildFast> (дата обращения: 23.05.2015).
9. Engle R.F., Granger C.W.J. Co-integration and error correction: Representation, estimation and testing // Econometrica. 1987. No. 55 (2). pp. 251–276.
10. FreeNAS 10 [Электронный ресурс] // Bug Tracking System: [сайт]. URL: <https://bugs.pcbsd.org/projects/freenas-10> (дата обращения: 30.08.2015).
11. FreeNAS 9 [Электронный ресурс] // Bug Tracking System: [сайт]. URL: <https://bugs.pcbsd.org/projects/freenas> (дата обращения: 30.08.2015).
12. Giove S. Fuzzy logic and Clustering methods for time series analisys // 2009 Internati onal Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy.
13. Goodwin P., Lawton R. On the asymmetry of the symmetric MAPE // International Journal of Forecasting. 1999. Vol. 15(4). pp. 405-408.
14. Gooijer J.G.D., Hyndman R.J. 25 years of time series forecasting // International Journal of Forecasting. 2006. Vol. 22(3). pp. 443-473.
15. Granger C. Some Properties of Time Series Data and Their Use in Econometric Model Specification // Journal of Econometrics. 1981. No. 16 (1). pp. 121-130.
16. Graves D., Pedrycz W. Multivariate Segmntntation of Timr Series with Differential Evolution // 2009 International Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy Logic and.

17. Gregory A.W., Hansen B.E. Residual-based tests for cointegration in models with regime shifts // *Journal of Econometrics*. 1996. No. 70 (1). pp. 99–126.
18. Herbst G., Bocklish S.F. Online Recognition of fuzzy time series patterns // 2009 International Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy.
19. Holt C.C. Forecasting trends and seasonals by exponentially weighted moving averages // O.N.R. Memorandum, Carnegie Inst. of Technology. 1997. Vol. 2.
20. Huarng K. Effective lengths of intervals to improve forecasting in fuzzy time series // *Fuzzy Sets and Systems*. 2001. Vol. 123. pp. 387–394.
21. Hyndman R.J., Koehler A.B. Another look at measures of forecast accuracy // *International Journal of Forecasting*. 2006. Vol. 22(4). pp. 679-688.
22. JIRA [Электронный ресурс] // Startpack : [сайт]. URL: <https://startpack.ru/application/jira> (дата обращения: 20.05.2015).
23. Johansen S. Cointegration and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models // *Econometrica*,. 1991. Vol. 59(6). pp. 1551–1580.
24. Kacprzyk J., Wilbik A. Using Fuzzy Linguistic summaries for the comparison of time series // 2009 International Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy Logic.
25. Koehler A.B. The asymmetry of the sAPE measure and other comments on the M3 - competition // *International Journal of Forecasting*. 2001. Vol. 17. pp. 570–574.
26. Kolassa S., Martin R. Percentage errors can ruin your day (and rolling the dice shows how) // *Foresight*. 2011. pp. 21-27.
27. Makridakis S. Accuracy measures: Theoretical and practical concerns // *International Journal of Forecasting*. 1993. Vol. 9. pp. 527–529.
28. Microsoft Project Server 2013 [Электронный ресурс] // Startpack: [сайт]. URL: <https://startpack.ru/application/microsoft-project-server> (дата обращения: 20.05.2015).
29. NN3 // Neural-forecasting-competition. URL: http://www.neural-forecasting-competition.com/downloads/NN3/datasets/NN3_REDUCED.xls (дата обращения: 20.01.2014).
30. OpenPanel [Электронный ресурс] // OpenPanel bugtracker: [сайт]. URL: <http://bugtracker.openpanel.com/projects/openpanel> (дата обращения: 30.08.2015).
31. PCBSD [Электронный ресурс] // Bug Tracking System: [сайт]. URL: <https://bugs.pcbsd.org/projects/pcbsd> (дата обращения: 30.08.2015).
32. Pedrycz W., Chen S.M. Time Series Analysis, Modeling and Applications: A Computational Intelligence Perspective (e-book Google) // *Intelligent Systems Reference Library*. 2013. Vol. 47. P. 404.
33. Perfilieva I., Novak V., Dvorak A. Fuzzy transform in the analysis of data // *Intern. Journal of Appr. reasoning*. 2008.
34. Perfilieva I., Dankov'a M., Bede B. Towards a Higher Degree F-Transform //

- Fuzzy Sets and Systems. 2011. Vol. 180. P. 319.
35. Perfilieva I., Yarushkina N., Afanaseva T. Relaxed Discrete F-Transform and its Application to the Time Series Analysis // Da Ruanetal (Eds.): Computational Intelligence. Foundations and Applications (Proc.of the 9th Int. FLINS Conf.). August 2010. pp. 249-255.
 36. Perfilieva I., Yarushkina N., Afanasieva T., Igonin A., Romanov A., Shishkina V. Soft computing tools for time series analysis and forecast // Proceedings of the 9th Int. Conf. on Application of Fuzzy Systems and Soft Computing (ICAFS 2010). August 2010. pp. 50-60.
 37. Perfilieva I., Yarushkina N., Afanasieva T., Romanov A. Time series analysis using soft computing methods // International journal of general systems. 2010. Vol. 42(6). pp. 687-705.
 38. Perfilieva I.G., Yarushkina N.G., Afanasieva T.V., Romanov A.A. Internet service for the analysis of enterprise economics using time series fuzzy modeling // Proceedings of the 2013 joint IFSA world congress and nafips annual meeting, IFSA/NAFIPS 2013. июнь 2013. pp. 1113-1118.
 39. Perfilieva I.G., Yarushkina N.G., Afanasieva T.V. Time series analysis by discrete f-transform // 2010 IEEE World Congress on Computational Intelligence, WCCI 2010 2010 6th IEEE World Congress on Computational Intelligence, WCCI 2010. July 2010. pp. 1-4.
 40. Phillips P.C.B., Ouliaris S. Asymptotic Properties of Residual Based Tests for Cointegration // Econometrica. Vol. 58(1). pp. 165–193.
 41. Romanov A., Yarushkina N., Perfilieva I. Time series grouping on the basis of f1-transform // IEEE international conference on fuzzy systems. июль 2014. pp. 517-521.
 42. Ruey-Chyn Tsaur J.C.O.Y.H.F.W. Fuzzy Relation Analysis in Fuzzy Time Series Model 2005. Vol. 49.
 43. Sabic D.A., Pedrycz W. Evaluation on fuzzy linear regression models // Fuzzy Sets and Systems. 1991. Vol. 23. pp. 51-63.
 44. Search for Issue [Электронный ресурс] // mongoDB: [сайт]. URL: <https://jira.mongodb.org/browse/WT-2271?jql=> (дата обращения: 20.11.2015).
 45. Shcherbakov M.V., Brebels A. Outliers and anomalies detection based on neural networks forecast procedure // Proceedings of the 31st Annual International Symposium on Forecasting (ISF 2011). 2011. pp. 21-22.
 46. Song Q., Chissom B. Forecasting enrollments with fuzzy time series – Part I // Fuzzy Sets and Systems. 1993. Vol. 54. pp. 1-9.
 47. Song Q., Chissom B. Fuzzy time series and its models 1993. Vol. 54. pp. 269-277.
 48. Tanaka H., Uejima S., Asai K. Linear regression analysis with fuzzy model // IEEE Transactions on Systems, Man and Cybernetics. 1982. No. 12 (6). pp. 903–907.
 49. Timina I.A. Time series clusterization // Interactive systems and technologies: the problem of Human-Computer Interaction (Collection of scientific papers). 2011. pp. 390-393.

50. Yarushkina N., Afanasieva T., Igonin A., Romanov A., Shishkina V., Perfilieva I. Time Series Processing and Forecasting using Soft Computing Tools // Lecture Notes in Computer Science (Proceedings of 13-th International Conf. RSFDGrC-2011). 2011. Vol. 6743 (XIII). pp. 155-163.
51. Yarushkina N., Afanasieva T., Timina I. Predicative analytics for developing software // SYRCoSE 2014. The 8th Spring / Summer Young Researchers' Colloquium on Software Engineering. May 2014. pp. 154-158.
52. Yu F., Pedrycz W., Yuan J. Finding Fuzzy Rules from Granular Time series // IFSA'05 FuzzyLogic, Soft Computing and Computational Intelligence.
53. Zadeh A.L. Fuzzy Sets. Information and Control, 1965.
54. Zadeh L.A., Gupta M.M., Ragade R.K., Yager R.R. Advances in Fuzzy Sets Theory and Applications // North-Holland Publishing Company. 1979. pp. 3-20.
55. Zadeh L.A. Generalized theory of uncertainty (GTU) – principal concepts and ideas // Computational statistic & Data analysis. 2006. Vol. 51. pp. 15-46.
56. Zadeh L.A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic // Fuzzy Sets and Systems. 1997. Vol. 90. pp. 111-127.
57. Алферов А. Методы точного прогнозирования поступлений денежных средств // Корпоративный менеджмент. 2013. URL: http://www.cfin.ru/finanalysis/math/cash_inflow.shtml (дата обращения: 23.05.2015).
58. Артюхов М.В., Ярушкина Н.Г., Романов А.А., Гуськов Г.Ю., Тимина И.А. Построение массива методов для прогнозирования временных рядов // VIII-ая Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте. май 2015. Т. 1. С. 332-341.
59. Артюхов М.В., Ярушкина Н.Г., Романов А.А., Гуськов Г.Ю., Тимина И.А. Прогнозирование временных рядов коллективами методов // Радиотехника. 2015. Т. 6. С. 48-54.
60. Афанасьева Т.В., Гуськов Г.Ю., Перфильева И.Г., Романов А.А., Тимина И.А., Ярушкина Н.Г. Интеллектуальный веб-сервис экспресс-анализа экономического состояния предприятия // Четырнадцатая национальная конференция по искусственному интеллекту с международным участием (КИИ-2014). 2014. Т. 2. С. 213-221.
61. Афанасьева Т.В., Наместников А.М., Перфильева И.Г., Романов А.А., Ярушкина Н.Г. Прогнозирование временных рядов: нечеткие модели. Ульяновск: УлГТУ, 2014. 145 с.
62. Афанасьева Т.В., Тимина И.А. Интеллектуальный анализ нечетких тенденций временных рядов // Сборник научных трудов VII-ой Международной научно-практической конференции «Интегрированные модели и мягкие вычисления. 2013. Т. 3. С. 474-483.
63. Афанасьева Т.В., Ярушкина Н.Г. Анализ эффективности модели нечеткой тенденции в прогнозировании временных рядов // Автоматизация процессов

- управления. 2011. Т. 4. С. 43-49.
64. Афанасьева Т.В., Ярушкина Н.Г. Нечеткий динамический процесс с нечеткими тенденциями в анализе временных рядов // Вестник Ростовского государственного университета путей сообщения. 2001. № 3 (43). С. 7-16.
 65. Афанасьева Т.В., Ярушкина Н.Г. Нечеткое моделирование временных рядов и анализ нечетких тенденций. Ульяновск: УлГТУ, 2009. 299 с.
 66. Афанасьева Т.В. Методология, модели и комплексы программ анализа временных рядов на основе нечетких тенденций, УлГТУ, Ульяновск, дисс. докт.техн.наук по спец.:05.13.18 2012.
 67. Афанасьева Т.В. Моделирование нечетких тенденций временных рядов. Ульяновск: УлГТУ, 2013. 215 с.
 68. Афанасьева Т.В. Модель ас1-шкалы для генерации лингвистических оценок в принятии решений // Вопросы современной науки и практики. университет им. в.и. вернадского. 2008. Т. 2(4). С. 91-96.
 69. Афанасьева Т.В. Нечеткое моделирование временных рядов и анализ нечетких тенденций. Ульяновск: УлГТУ, 2009. 299 с.
 70. Афанасьева Т.В. Прогнозирование локальных тенденций временных рядов в задачах анализа больших данных // Автоматизация процессов управления. 2014. Т. 4(38). С. 42-47.
 71. Афанасьева Т.В. Решение задач интеллектуального анализа временных рядов в рамках структурно-лингвистического подхода // Автоматизация процессов управления. 2010. Т. 2. С. 54-58.
 72. Батыршин И.З., Недосекин А.О., Стецко А.А., Тарасов В.Б., Язенин А.В. Нечеткие гибридные системы. Теория и практика. М.: ООО Издательская фирма "Физико-математическая литература", 2007. 208 с.
 73. Бокс Д., Дженкинс Г. Анализ временных рядов. Прогноз и управление. М: Мир, 1974. 406 с.
 74. Боэм Б., Браун Д., Каспар Х., Липоу М. Характеристики качества программного обеспечения. М.: Мир, 1981. 208 с.
 75. Воронина В.В., Тимина И.А. Разработка модуля анализа многомерных временных рядов для системы оценки финансового состояния предприятия // Сборник научных трудов 1-ой всероссийской научно-практической конференции "Прикладные информационные системы". 2015. С. 4-10.
 76. Воронина В.В. Математическое моделирование и численное исследование поведения агрегатов летательных аппаратов в задачах диагностики на основе гранулированных временных рядов, УлГТУ, Ульяновск, диссертация 2011. 171 с.
 77. Временной ряд [Электронный ресурс] // MachineLearning.ru: [сайт]. URL: http://www.machinelearning.ru/wiki/index.php?title=Временной_ряд (дата обращения: 05.01.2014).
 78. ГОСТ Р 50.1.028-2001 Информационные технологии поддержки жизненного цикла продукции. Методология функционального моделирования. М.:

ГОССТАНДАРТ РОССИИ, 2001.

79. Зайдель А.Н. Элементарные оценки ошибок измерений. 2-е изд. испр. и доп.-е изд. Л.: «Наука», 1967. 88 с.
80. Ивантер Э.В., Коросов А.В. Основы биометрии: введение в статистический анализ биологических явлений и процессов. Петрозаводск: Изд-во ПГУ, 1992. 164 с.
81. Кендэл М. Временные ряды. М.: Финансы и статистика, 1981. 199 с.
82. Климова А.С. Модели и методы гибридной реляционной кластеризации данных, Казань, Диссертация 2013. 107 с.
83. Коновалов Ю.Е., Тронин В.Г. Распределение задач по развитию информационной системы и внедрению новых информационных технологий НПО // Автоматизация процессов управления. 2006. Т. 2(8). С. 204-208.
84. Кремер Н.Ш., Путков Б.А. Эконометрика: учебник для вузов. М.: ЮНИТИ-ДАНА, 2002. 311 с.
85. Кроль Т.Я., Харин М.А. Методы решения задачи кластеризации и прогнозирования в электронном архиве // Молодой ученый. 2011. Т. 1. № 6. С. 135-137.
86. Круглов В.В., Дли М.И., Голунов Р.Ю. Нечеткая логика и искусственные нейронные сети. М.: Физматлит, 2001. 224 с.
87. Меламед М.А. Современные методы анализа и прогнозирования режимов электропотребления в электроэнергетических системах // Итоги науки и техники. – Сер. Энергетические системы и их автоматизация. 1988. Т. 4. С. 4-11.
88. Методические рекомендации по управлению проектами в области информационных технологий [Электронный ресурс] // Pandia: [сайт]. [2011]. URL: <http://pandia.ru/text/78/208/45555.php> (дата обращения: 15.03.2015).
89. Мидоу Ч. Анализ информационно-поисковых систем. М.: Мир, 1970.
90. Наместников А.М. Интеллектуальные проектные репозитории. Ульяновск: УлГТУ, 2009. 111 с.
91. Островский А.А. Нечеткая кластеризация электронных информационных ресурсов проектного репозитория при автоматизированном проектировании, УлГТУ, Ульяновск, дис. канд.техн.наук: спец. 05.13.12 2010.
92. Перфильева И.Г., Романов А.А. Извлечение знаний о группах временных рядов для задач прогнозирования // Четырнадцатая национальная конференция по искусственному интеллекту с международным участием кии-2014. сентябрь 2014. С. 75-83.
93. Перфильева И.Г., Ярушкина Н.Г., Романов А.А. Моделирование нестационарных временных рядов // Гибридные и синергетические интеллектуальные системы. июль 2014. С. 281-289.
94. Перфильева И.Г., Ярушкина Н.Г., Романов А.А. Формирование групп временных рядов // VI-я Всероссийская научно-практическая конференция

- Нечеткие системы и мягкие вычисления–2014 (НСМВ–2014). июнь 2014. Т. 1. С. 8-16.
95. Поль М.Д., Стивен М.М.И., Эндрю Г. Непрерывная интеграция: улучшение качества программного обеспечения и снижение риска. Вильямс, 2008. 240 с.
 96. Предиктор [Электронный ресурс] // Академик: [сайт]. URL: <http://dic.academic.ru/dic.nsf/ruwiki/302122/Предиктор> (дата обращения: 15.06.2015).
 97. Программное обеспечение для управления проектами [Электронный ресурс] // Википедия: [сайт]. URL: https://ru.wikipedia.org/wiki/Программное_обеспечение_для_управления_проектами (дата обращения: 20.05.2015).
 98. Радионова Ю.А. Интеллектуальные методы организации архивов технической документации научно-производственного объединения, УлГТУ, Ульяновск, диссертация 2012. 302 с.
 99. Регрессия, корреляция и совпадение [Электронный ресурс] // IEEE: [сайт]. URL: <http://ieee.tpu.ru/system/regression.html> (дата обращения: 20.05.2015).
 100. Романов А.А. Математическое моделирование и комплекс программ анализа временных рядов на основе нечеткой модели, УлГТУ, Ульяновск, диссертация 2013. 147 с.
 101. Романов А.А. Моделирование и прогнозирование временных рядов на основе метода F-преобразования // Автоматизация процессов управления. 2012. Т. 2. С. 28-31.
 102. Романов А.А. Применение метода F-преобразования для прогноза тренда и числового представления временного ряда // Известия Самарского научного центра Российской академии наук. 2011. Т. 13. С. 1103-1109.
 103. Саркисян С.А., Каспин В.И., Лисичкин В.А., Минаев Э.С., Пасечник Г.С. Теория прогнозирования и принятия решений. М.: Высшая школа, 1977. 351 с.
 104. Свод знаний по управлению проектами [Электронный ресурс] // Википедия: [сайт]. URL: http://gruzdoff.ru/wiki/Свод_знаний_по_управлению_проектами (дата обращения: 20.05.2015).
 105. Система отслеживания ошибок [Электронный ресурс] // Википедия: [сайт]. URL: https://ru.wikipedia.org/wiki/Система_отслеживания_ошибок (дата обращения: 20.05.2015).
 106. Система управления версиями [Электронный ресурс] // Википедия: [сайт]. URL: https://ru.wikipedia.org/wiki/Система_управления_версиями (дата обращения: 20.05.2015).
 107. Тимина И.А., Радионова Ю.А., Эгов Е.Н. Мера энтропии в решениях задач управления архива конструкторской технологической документации // VIII-ая Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте. май 2015. Т. 1. С. 361-370.
 108. Тимина И.А., Радионова Ю.А., Ярушкина Н.Г., Эгов Е.Н. Анализ временных

- рядов в задачах управления архива конструкторской технологической документации // Радиотехника. 2015. Т. 6. С. 83-88.
109. Тимина И.А., Радионова Ю.А. Предикативное прогнозирование поступления электронной технической документации в проектном репозитории на основе нечетких тенденций // Сборник материалов 49й Научно-технической конференции «Вузовская наука в современных условиях». 2015. С. 194-197.
 110. Тимина И.А. Анализ и прогнозирование многомерных временных рядов // Научная сессия НИЯУ МИФИ-2013. 2013. Т. 2. С. 318.
 111. Тимина И.А. К вопросу о степени сходства нечетких тенденций в интеллектуальном анализе временных рядов // Научная сессия НИЯУ МИФИ-2014. 2014. Т. 3. С. 120.
 112. Тимина И.А. Кластеризация временных рядов // Информатика и вычислительная техника. 2012. С. 292-298.
 113. Тимина И.А. Кластеризация многомерных временных рядов // Информатика, моделирование, автоматизация проектирования. 2012. С. 374-378.
 114. Тимина И.А. Корректировка гипотез прогноза для извлечения знаний о временных рядах // Четырнадцатая национальная конференция по искусственному интеллекту с международным участием (КИИ-2014). 2014. Т. 3. С. 68-76.
 115. Тимина И.А. Нечеткая зависимость временных рядов в моделировании и прогнозировании состояния организации // Информатика и вычислительная техника (сборник научных трудов 5-й Всероссийской научно-технической. Май 2013. С. 228-234.
 116. Тимина И.А. Нечеткая зависимость как метод решения задач интеллектуального анализа временных рядов // Автоматизация процессов управления. 2013. Т. 3(33). С. 39-44.
 117. Тимина И.А. Прогнозирование зависимых нестационарных временных рядов на основе нечетких элементарных тенденций с учетом регулирования гипотезы // Сборник материалов 48й Научно-технической конференции «Вузовская наука в современных условиях». 2014. С. 182-185.
 118. Тимина И.А. Прогнозирование нечетких временных рядов на основе гипотез о поведении тенденций // Сборник материалов 47й Научно-технической конференции «Вузовская наука в современных условиях». 2013. С. 262-265.
 119. Тренд [Электронный ресурс] // MachineLearning.ru : [сайт]. URL: <http://www.machinelearning.ru/wiki/index.php?title=Тренд> (дата обращения: 05.01.2014).
 120. Управление проектами [Электронный ресурс] // Википедия: [сайт]. URL: https://ru.wikipedia.org/wiki/Управление_проектами (дата обращения: 20.05.2015).
 121. Чекина А.В. Генетическая кластеризация технической документации в проектном репозитории САПР, УлГТУ, Ульяновск, диссертация 2012. 215 с.
 122. Чубукова И.А. Data Mining: учебный курс. Интернет-университет

- информационных технологий, Бином, 2008. 324 с.
123. Щербаков М.В., Бребельс А., Щербакова Н.Л., Тюков А.П. Обзор оценок качества моделей прогнозирования URL: http://www.mtas.ru/bitrix/components/bitrix/forum.interface/show_file.php?fid=6450 (дата обращения: 25.03.2015).
 124. Экономические коэффициенты [Электронный ресурс] // Система анализа экономических показателей предприятий: [сайт]. URL: <http://tsas.ulstu.ru/> (дата обращения: 15.12.2012).
 125. Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов: уч.пособие. Ульяновск: УлГТУ, 2010. 320 с.
 126. Ярушкина Н.Г., Афанасьева Т.В., Перфильева И.Г. Интеллектуальный анализ временных рядов. М.: ИД "ФОРМУ": ИНФРА-М, 2012. 160 с.
 127. Ярушкина Н.Г., Афанасьева Т.В., Романов А.А., Тимина И.А. Извлечение знаний о зависимостях временных рядов для задач прогнозирования // Радиотехника. 2014. Т. 7. С. 141-146.
 128. Ярушкина Н.Г., Афанасьева Т.В., Тимина И.А. К вопросу о прогнозировании многомерного гетерогенного временного ряда // Тринадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2012. 2012. Т. 1. С. 149-155.
 129. Ярушкина Н.Г., Афанасьева Т.В., Тимина И.А. Нечеткая грануляция в моделировании и прогнозировании объема телекоммуникационного трафика // Научные технологии. 2013. Т. 14(5). С. 67-73.
 130. Ярушкина Н.Г., Афанасьева Т.В. Интеллектуальный анализ поведения сложных систем // Радиотехника. 2012. Т. 9. С. 10-13.
 131. Ярушкина Н.Г., Афанасьева Т.В. Математическое моделирование коротких временных рядов на основе нечетких тенденций // Обзор прикладной и промышленной математики. 2011. Т. 18(4). С. 564-570.
 132. Ярушкина Н.Г., Воронина В.В., Тимина И.А., Эгов Е.Н. Прогнозирование состояния технической системы с применением меры энтропии для нечетких временных рядов // Автоматизация процессов управления. 2015. Т. 3(41). С. 49-58.
 133. Ярушкина Н.Г., Наместников А.М. Методы мягких вычислений в организации хранилища проектных документов // Радиотехника. 2012. Т. 9. С. 14-20.
 134. Ярушкина Н.Г., Перфильева И.Г., Афанасьева Т.В. Интегральный метод нечеткого моделирования и анализа нечетких тенденций // Автоматизация процессов управления. 2010. Т. 2. С. 59-63.
 135. Ярушкина Н.Г., Перфильева И.Г., Афанасьева Т.В. Интеграция нечетких моделей для анализа временных рядов // Известия самарского научного центра российской академии наук. 2010. Т. 12(4-2). С. 506-509.
 136. Ярушкина Н.Г., Перфильева И.Г., Игонин А.Г., Романов А.А., Юнусов Т.Р., Шишкина В.В. Разработка internet-сервиса, интегрирующего нечеткое

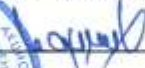
- моделирование и анализ нечетких тенденций временных рядов // Автоматизация процессов управления. 2010. Т. 2. С. 64-69.
137. Ярушкина Н.Г., Тимина И.А. Алгоритм адаптации прогноза зависимых временных рядов // 2-ой Международный Пospelовский симпозиум «Гибридные и синергетические интеллектуальные системы ГИСИС'2014». 2014. С. 398-403.
138. Ярушкина Н.Г., Тимина И.А. Коэффициент подобия как способ оценки влияния наличия нечетких тенденций шаблона на прогноз // VI-я Всероссийская научно-практическая конференция Нечеткие системы и мягкие вычисления–2014 (НСМВ–2014). 2014. С. 90-98.
139. Ярушкина Н.Г., Тимина И.А. Модель и средства управления проектированием автоматизированной системы на основе динамики метрик программного кода // Автоматизация процессов управления. 2015. Т. 3(41). С. 73-81.
140. Ярушкина Н.Г., Чекина А.В. Кластеризация информационных ресурсов на основе генетического алгоритма // Автоматизация процессов управления. 2010. Т. 4. С. 66-70.
141. Ярушкина Н.Г. Автоматизированное проектирование сложных технических систем в условиях неопределенности, УлГТУ, Ульяновск, дисс. докт.техн.наук по спец.: 05.13.12 1998.
142. Ярушкина Н.Г. Интеллектуализация автоматизированного проектирования сложных технических систем в условиях неопределенности // Автоматизация процессов управления. 2011. Т. 1. С. 13-19.
143. Ярушкина Н.Г. Методы нечетких экспертных систем в интеллектуальных САПР. Саратов: Изд-во Саратов.ун-та, 1997. 107 с.
144. Ярушкина Н.Г. Нечеткие системы: обзор итогов и тенденций развития // Искусственный интеллект и принятие решений. 2008. Т. 4. С. 26-38.
145. Ярушкина Н.Г. Основы теории нечетких и гибридных систем. М.: "Финансы и статистика", 2004. 320 с.

ПРИЛОЖЕНИЕ А. АКТ ВНЕДРЕНИЯ ПСАП СКВ В ФНПЦ АО «НПО «МАРС» (Г. УЛЬЯНОВСК)

УТВЕРЖДАЮ

Генеральный директор,
председатель НТС ФНПЦ

АО «НПО «Марс», к.т.н.

 В.А. Маклаев

«» 2015 г.

А К Т

об использовании результатов кандидатской диссертации И.А. Тиминой
“Автоматизированное проектирование на основе модели прогнозирования нечетких
тенденций метрик проектной деятельности”

Научно-техническая комиссия в составе:

председателя первый заместитель генерального директора по науке, к.т.н.

комиссии: Э.Д. Павлыгин,

членов заместитель главного инженера по качеству и инженерно-

комиссии: техническому обеспечению – начальник управления 5, к.т.н.

А.А. Емельянов,

начальник отдела технической документации А.Е. Ефремов,

ведущий инженер-программист, к.т.н. Ю.А. Радионова.

настоящим актом подтверждает использование следующих научных и практических результатов диссертационной работы И.А. Тиминой “Автоматизированное проектирование на основе модели прогнозирования нечетких тенденций метрик проектной деятельности” для анализа статистики событий системы контроля версий различных проектов и статистики деятельности отдела технической документации предприятия:

- алгоритм прогнозирования тенденций временных рядов на основе нечеткого моделирования с использованием системы нечетких правил для различных гипотез прогноза;
- алгоритм прогнозирования тенденций временных рядов на основе нечетких мер сходства временных рядов, отличающихся предложенными индикаторами сходства и процедурами их построения и внесения поправок в прогноз;
- комплекс программ моделирования как подсистема автоматизации системы контроля версий (СКВ) на основе анализа и прогнозирования динамики метрик

проектной деятельности.

Разработанная программная система анализа динамики метрик проектной деятельности используется в работе отдела технической документации.

Эффективность использования научно-технических результатов подтверждена экспериментальными исследованиями, целью которых являлось построение продуктивного средства управления совокупностью проектов крупной проектной организации на основе анализа и моделирования состояния метрик проектов программно-аппаратных комплексов.

Для реализации автоматизации системы контроля версий в электронном архиве на ФНПЦ АО «НПО «Марс» была разработана формализованная модель состояния проектной деятельности на основе метрик проектов, загруженных в систему контроля версий проектного репозитория; средство анализа состояния метрик сущностей СКВ проектов программного обеспечения как инструмент управления работами организации; алгоритмы прогнозирования нечетких временных рядов метрических характеристик проекта. Проведенные вычислительные эксперименты по исследованию эффективности предложенных методов и алгоритмов показали, что больше, чем в половине случаев предложенный подход показывает улучшение качества прогнозирования по сравнению с моделями класса *ARIMA* по критерию *MSE*, который составляет не более 1-2% от среднего значения метрик прогноза.

Председатель комиссии:

Первый заместитель генерального директора
по науке, к.т.н.



Э.Д. Павлыгин

Члены комиссии:

Заместитель главного инженера
по качеству и инженерно-техническому
обеспечению – начальник управления 5, к.т.н.



А.А. Емельянов

Начальник отдела
технической документации



А.Е. Ефремов

Ведущий инженер-программист, к.т.н.



Ю.А. Радионова

**ПРИЛОЖЕНИЕ В. АКТ СДАЧИ-ПРИЕМКИ НА ВЫПОЛНЕНИЕ РАБОТЫ
«РАЗРАБОТКА НЕЧЕТКИХ МОДЕЛЕЙ И РЕАЛИЗАЦИЯ НЕЧЕТКИХ
МЕТОДОВ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ» ООО
«ЭВЕРЕСТ РЕСЕРЧ» (Г. МОСКВА)**

Наименование исполнителя:

федеральное государственное бюджетное
образовательное учреждение высшего
профессионального образования «Ульяновский
государственный технический университет»
432027, г. Ульяновск, ул. Северный Венец, д. 32
ИНН 7325000052/ КПП 732501001
УФК по Ульяновской области
(Ульяновский государственный технический
университет д/сч 20686Х85090)
Банк: ГРКЦ ГУ Банка России по Ульяновской
обл., г. Ульяновск,
БИК 047308001
р/сч 40503810900001000001
ОКПО 02069378 ОКОНХ 92110
Назначение платежа:
Доходы от выполнения
научно-исследовательских и опытно-
конструкторских работ
ОФК 6800 (07330201010010000130)

Наименование заказчика:

ООО «Эверест Ресерч»
Адрес: 115432, г. Москва,
ул. Бориса Галушкина, д. 14/1
Р/сч: 40702810802610000143 в
ОАО «Альфа-Банк» г. Москва
К/сч: 30101810200000000593
Банк: ОАО «Альфа-Банк» г.
Москва
ИНН: 7728168971
КПП: 771701001

АКТ

сдачи-приемки на выполненные работы

«Разработка нечетких моделей и реализация нечетких методов прогнозирования
временных рядов»

по этапу договора № ДЗ46 от 23 сентября 2014г.

Составлен 20 марта 2015г.

Мы, нижеподписавшиеся, представитель Исполнителя - проректор по НР Государственного образовательного учреждения высшего профессионального образования «Ульяновский Государственный Технический Университет» Ярушкина Н.Г., с одной стороны, и представитель Заказчика – генеральный директор ООО «Эверест Ресерч» Артюхов М.В., с другой стороны, составили настоящий акт о том, что разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов выполнены в полном объеме и удовлетворяют условиям договора.

Право собственности у Заказчика возникает с момента передачи ему результатов выполненной работы в целом и оплаты им выполненных работ. Объем выполненных работ составляет 200 000 (Двести тысяч) рублей. НДС не предусмотрен.

Работу сдал
от Исполнителя



Н.Г. Ярушкина

Работу принял
от Заказчика



М.В. Артюхов

Дополнительное соглашение №1 от «23» декабря 2014г.
к договору № Д346/31 от 23 сентября 2014г.
«Разработка нечетких моделей и реализация нечетких методов прогнозирования
временных рядов»

«23» декабря 2014г.

г. Ульяновск

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Ульяновский государственный технический университет» (сокращенное наименование — Ульяновский государственный технический университет, УлГТУ), именуемое в дальнейшем «Исполнитель», в лице первого проректора, проректора по научной работе Ярушкиной Н.Г., действующего на основании доверенности № 41 от 01.04.2014, с одной стороны, и ООО «Эверест Ресерч», именуемое в дальнейшем «Заказчик», в лице Артюхова Михаила Васильевича, действующего на основании Устава, с другой стороны, заключили настоящее дополнительное соглашение №1 от 23.12.2014г. к договору № Д346/31 от 23.09.2014г. «Разработка нечетких моделей и реализация нечетких методов прогнозирования временных рядов» о нижеследующем:

1. Считать сроком окончания договора 30.04.2015г.:
2. Настоящее Дополнительное соглашение №1 является неотъемлемой частью договора № Д346/31 от 23.09.2014г., составлено в 2-х экземплярах по одному для каждой из Сторон и вступает в силу с момента его подписания Сторонами.

ЗАКАЗЧИК
От ООО «Эверест Ресерч»
Генеральный директор
М.В. Артюхов



ИСПОЛНИТЕЛЬ
От Ульяновского государственного
технического университета
Первый проректор,
проректор по научной работе
Н.Г. Ярушкина



**ПРИЛОЖЕНИЕ С. СВИДЕТЕЛЬСТВО О ГОСУДАРСТВЕННОЙ
РЕГИСТРАЦИИ ПРОГРАММЫ ДЛЯ ЭВМ**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО
о государственной регистрации программы для ЭВМ
№ 2014612213

**Интеллектуальный экспресс-анализатор сходства нечетких
тенденций временных рядов**

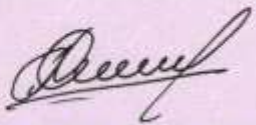
Правообладатель: *федеральное государственное бюджетное
образовательное учреждение высшего профессионального
образования «Ульяновский государственный технический
университет» (RU)*

Авторы: *Ярушкина Надежда Глебовна (RU),
Тимина Ирина Александровна (RU)*

Заявка № **2013662131**
Дата поступления **25 декабря 2013 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **21 февраля 2014 г.**



Руководитель Федеральной службы
по интеллектуальной собственности


Б.П. Симонов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2015617046

**Интегрированная система нечетких методов
прогнозирования временных рядов**

Правообладатель: *федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Ульяновский государственный технический университет» (RU)*

Авторы: *Ярушкина Надежда Глебовна (RU), Романов Антон Алексеевич (RU), Гуськов Глеб Юрьевич (RU), Тимина Ирина Александровна (RU)*



Заявка № **2015613884**

Дата поступления **12 мая 2015 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ **29 июня 2015 г.**

*Врио руководителя Федеральной службы
по интеллектуальной собственности*

Л.Л. Кирий

**ПРИЛОЖЕНИЕ D. СЕРТИФИКАТ ПОБЕДИТЕЛЯ МЕЖДУНАРОДНОГО
КОНКУРСА ПО ВРЕМЕННЫМ РЯДАМ «ВЫЧИСЛИТЕЛЬНЫЙ
ИНТЕЛЛЕКТ В ПРОГНОЗИРОВАНИИ»**



ПРИЛОЖЕНИЕ Е. ДАННЫЕ ПОСТУПЛЕНИЯ ОТЧЕТОВ ПО ИЗДЕЛИЯМ ФНПЦ АО «НПО «МАРС»

Таблица Е.1 – Данные по изделию 142

Значение	Код изделия	Дата, время
2	142	10.09.1909 0:00
51	142	30.09.1996 0:00
1	142	13.01.1997 0:00
33	142	23.01.1997 0:00
9	142	04.03.1997 0:00
84	142	13.10.1998 0:00
2	142	11.01.2000 0:00
5	142	18.01.2000 0:00
2	142	23.02.2000 0:00
5	142	29.02.2000 0:00
1	142	01.03.2000 0:00
4	142	09.03.2000 0:00
1	142	14.03.2000 0:00
7	142	29.03.2000 0:00
7	142	07.04.2000 0:00
6	142	14.04.2000 0:00
2	142	15.05.2000 0:00
1	142	16.05.2000 0:00
2	142	19.05.2000 0:00
6	142	24.05.2000 0:00
22	142	29.05.2000 0:00
3	142	27.06.2000 0:00
2	142	29.06.2000 0:00
2	142	04.07.2000 0:00
10	142	05.07.2000 0:00
5	142	17.07.2000 0:00
5	142	18.07.2000 0:00
28	142	20.07.2000 0:00
9	142	27.07.2000 0:00
1	142	28.07.2000 0:00
14	142	02.08.2000 0:00
3	142	21.08.2000 0:00
1	142	11.09.2000 0:00
1	142	29.09.2000 0:00
2	142	02.10.2000 0:00
2	142	11.10.2000 0:00
2	142	19.10.2000 0:00
4	142	04.11.2000 0:00
1	142	13.11.2000 0:00
1	142	19.01.2001 0:00
1	142	20.03.2001 0:00
5	142	27.03.2001 0:00
4	142	28.03.2001 0:00
2	142	13.04.2001 0:00
1	142	07.05.2001 0:00
12	142	21.06.2001 0:00
5	142	02.07.2001 0:00
4	142	03.07.2001 0:00
2	142	13.07.2001 0:00
2	142	15.08.2001 0:00
7	142	28.08.2001 0:00
1	142	26.09.2001 0:00
1	142	01.10.2001 0:00
4	142	04.10.2001 0:00
2	142	01.11.2001 0:00
1	142	02.11.2001 0:00
1	142	14.11.2001 0:00
1	142	06.12.2001 0:00
2	142	17.12.2001 0:00
3	142	25.12.2001 0:00
9	142	01.02.2002 0:00
1	142	12.02.2002 0:00
3	142	14.03.2002 0:00
2	142	21.03.2002 0:00
2	142	28.03.2002 0:00
2	142	08.04.2002 0:00
1	142	10.04.2002 0:00
3	142	15.04.2002 0:00

2	142	17.04.2002 0:00
6	142	25.04.2002 0:00
1	142	13.06.2002 0:00
93	142	21.06.2002 0:00
3	142	24.06.2002 0:00
2	142	26.06.2002 0:00
52	142	27.06.2002 0:00
1	142	03.07.2002 0:00
1	142	08.07.2002 0:00
26	142	10.07.2002 0:00
1	142	18.07.2002 0:00
2	142	29.07.2002 0:00
3	142	31.07.2002 0:00
1	142	01.08.2002 0:00
2	142	14.08.2002 0:00
2	142	15.08.2002 0:00
1	142	20.08.2002 0:00
4	142	27.08.2002 0:00
1	142	10.09.2002 0:00
1	142	11.09.2002 0:00
4	142	16.09.2002 0:00
4	142	01.10.2002 0:00
2	142	08.10.2002 0:00
1	142	28.10.2002 0:00
1	142	30.10.2002 0:00
3	142	19.11.2002 0:00
3	142	25.11.2002 0:00
2	142	28.11.2002 0:00
3	142	04.12.2002 0:00
1	142	09.12.2002 0:00
1	142	10.12.2002 0:00
2	142	14.12.2002 0:00
2	142	17.12.2002 0:00
2	142	31.12.2002 0:00
2	142	13.01.2003 0:00
14	142	16.01.2003 0:00
1	142	25.01.2003 0:00
3	142	03.02.2003 0:00
2	142	04.02.2003 0:00
2	142	05.02.2003 0:00
1	142	06.02.2003 0:00
1	142	07.02.2003 0:00
5	142	08.02.2003 0:00
37	142	09.02.2003 0:00
2	142	12.02.2003 0:00
2	142	14.02.2003 0:00
2	142	20.02.2003 0:00
1	142	05.03.2003 0:00
2	142	14.03.2003 0:00
2	142	18.03.2003 0:00
1	142	20.03.2003 0:00
1	142	31.03.2003 0:00
3	142	02.04.2003 0:00
1	142	11.04.2003 0:00
2	142	14.04.2003 0:00
4	142	15.04.2003 0:00
1	142	29.04.2003 0:00
1	142	05.05.2003 0:00
1	142	22.05.2003 0:00
3	142	27.05.2003 0:00
3	142	28.05.2003 0:00
5	142	10.06.2003 0:00
1	142	16.06.2003 0:00
1	142	17.06.2003 0:00
1	142	18.06.2003 0:00
33	142	23.06.2003 0:00
1	142	01.07.2003 0:00
2	142	03.07.2003 0:00
2	142	08.07.2003 0:00
6	142	10.07.2003 0:00
3	142	11.07.2003 0:00
1	142	14.07.2003 0:00

4	142	15.07.2003 0:00
8	142	23.07.2003 0:00
2	142	30.07.2003 0:00
3	142	01.08.2003 0:00
1	142	04.08.2003 0:00
1	142	05.08.2003 0:00
6	142	07.08.2003 0:00
1	142	14.08.2003 0:00
2	142	18.08.2003 0:00
2	142	21.08.2003 0:00
1	142	22.08.2003 0:00
2	142	26.08.2003 0:00
1	142	29.08.2003 0:00
1	142	01.09.2003 0:00
1	142	02.09.2003 0:00
4	142	19.09.2003 0:00
1	142	22.09.2003 0:00
1	142	26.09.2003 0:00
6	142	29.09.2003 0:00
1	142	07.10.2003 0:00
1	142	16.10.2003 0:00
2	142	20.10.2003 0:00
5	142	23.10.2003 0:00
4	142	24.10.2003 0:00
2	142	25.10.2003 0:00
4	142	28.10.2003 0:00
11	142	29.10.2003 0:00
16	142	30.10.2003 0:00
2	142	06.11.2003 0:00
2	142	10.11.2003 0:00
2	142	11.11.2003 0:00
2	142	13.11.2003 0:00
2	142	21.11.2003 0:00
4	142	24.11.2003 0:00
3	142	25.11.2003 0:00
3	142	26.11.2003 0:00
1	142	29.11.2003 0:00
2	142	02.12.2003 0:00
3	142	08.12.2003 0:00
1	142	10.12.2003 0:00
1	142	11.12.2003 0:00
1	142	17.12.2003 0:00
1	142	18.12.2003 0:00
1	142	22.12.2003 0:00
3	142	27.12.2003 0:00
2	142	30.12.2003 0:00
3	142	31.12.2003 0:00
10	142	06.01.2004 0:00
7	142	13.01.2004 0:00
1	142	17.01.2004 0:00
1	142	19.01.2004 0:00
12	142	20.01.2004 0:00
5	142	21.01.2004 0:00
2	142	22.01.2004 0:00
7	142	26.01.2004 0:00
8	142	28.01.2004 0:00
4	142	30.01.2004 0:00
4	142	02.02.2004 0:00
1	142	03.02.2004 0:00
4	142	04.02.2004 0:00
1	142	07.02.2004 0:00
1	142	09.02.2004 0:00
1	142	12.02.2004 0:00
3	142	17.02.2004 0:00
8	142	26.02.2004 0:00
2	142	01.03.2004 0:00
4	142	04.03.2004 0:00
7	142	11.03.2004 0:00
4	142	12.03.2004 0:00
4	142	13.03.2004 0:00
1	142	15.03.2004 0:00
1	142	16.03.2004 0:00
3	142	18.03.2004 0:00
4	142	26.03.2004 0:00
5	142	29.03.2004 0:00
8	142	30.03.2004 0:00
8	142	02.04.2004 0:00
1	142	04.04.2004 0:00
3	142	07.04.2004 0:00

1	142	08.04.2004 0:00
5	142	13.04.2004 0:00
8	142	15.04.2004 0:00
1	142	20.04.2004 0:00
2	142	26.04.2004 0:00
3	142	29.04.2004 0:00
2	142	05.05.2004 0:00
3	142	13.05.2004 0:00
5	142	14.05.2004 0:00
2	142	17.05.2004 0:00
5	142	18.05.2004 0:00
1	142	19.05.2004 0:00
1	142	21.05.2004 0:00
3	142	31.05.2004 0:00
1	142	01.06.2004 0:00
1	142	02.06.2004 0:00
10	142	04.06.2004 0:00
2	142	10.06.2004 0:00
1	142	11.06.2004 0:00
1	142	16.06.2004 0:00
1	142	18.06.2004 0:00
1	142	21.06.2004 0:00
1	142	28.06.2004 0:00
1	142	30.06.2004 0:00
2	142	05.07.2004 0:00
2	142	06.07.2004 0:00
1	142	07.07.2004 0:00
6	142	12.07.2004 0:00
3	142	15.07.2004 0:00
1	142	19.07.2004 0:00
5	142	20.07.2004 0:00
2	142	29.07.2004 0:00
2	142	02.08.2004 0:00
4	142	03.08.2004 0:00
1	142	04.08.2004 0:00
4	142	05.08.2004 0:00
1	142	06.08.2004 0:00
4	142	09.08.2004 0:00
3	142	11.08.2004 0:00
2	142	12.08.2004 0:00
1	142	16.08.2004 0:00
1	142	18.08.2004 0:00
1	142	23.08.2004 0:00
2	142	25.08.2004 0:00
2	142	02.09.2004 0:00
8	142	06.09.2004 0:00
5	142	07.09.2004 0:00
8	142	08.09.2004 0:00
1	142	10.09.2004 0:00
1	142	13.09.2004 0:00
1	142	14.09.2004 0:00
5	142	23.09.2004 0:00
2	142	24.09.2004 0:00
2	142	27.09.2004 0:00
3	142	28.09.2004 0:00
2	142	30.09.2004 0:00
2	142	04.10.2004 0:00
2	142	05.10.2004 0:00
2	142	08.10.2004 0:00
2	142	18.10.2004 0:00
5	142	22.10.2004 0:00
1	142	27.10.2004 0:00
6	142	01.11.2004 0:00
2	142	06.11.2004 0:00
1	142	10.11.2004 0:00
5	142	11.11.2004 0:00
5	142	17.11.2004 0:00
2	142	19.11.2004 0:00
2	142	23.11.2004 0:00
3	142	24.11.2004 0:00
2	142	25.11.2004 0:00
1	142	30.11.2004 0:00
1	142	06.12.2004 0:00
1	142	07.12.2004 0:00
3	142	09.12.2004 0:00
1	142	11.12.2004 0:00
1	142	14.12.2004 0:00
2	142	16.12.2004 0:00
2	142	23.12.2004 0:00

3	142	24.12.2004 0:00
4	142	27.12.2004 0:00
4	142	29.12.2004 0:00
3	142	17.01.2005 0:00
8	142	21.01.2005 0:00
1	142	25.01.2005 0:00
5	142	31.01.2005 0:00
3	142	01.02.2005 0:00
1	142	02.02.2005 0:00
1	142	09.02.2005 0:00
7	142	16.02.2005 0:00
1	142	18.02.2005 0:00
1	142	19.02.2005 0:00
5	142	21.02.2005 0:00
2	142	26.02.2005 0:00
1	142	02.03.2005 0:00
1	142	11.03.2005 0:00
3	142	12.03.2005 0:00
4	142	15.03.2005 0:00
3	142	19.03.2005 0:00
5	142	22.03.2005 0:00
1	142	23.03.2005 0:00
2	142	24.03.2005 0:00
4	142	26.03.2005 0:00
9	142	28.03.2005 0:00
12	142	29.03.2005 0:00
1	142	30.03.2005 0:00
1	142	06.04.2005 0:00
5	142	07.04.2005 0:00
3	142	08.04.2005 0:00
5	142	12.04.2005 0:00
3	142	15.04.2005 0:00
1	142	18.04.2005 0:00
1	142	20.04.2005 0:00
1	142	21.04.2005 0:00
3	142	22.04.2005 0:00
2	142	26.04.2005 0:00
3	142	27.04.2005 0:00
1	142	29.04.2005 0:00
1	142	03.05.2005 0:00
10	142	14.05.2005 0:00
7	142	16.05.2005 0:00
2	142	18.05.2005 0:00
12	142	19.05.2005 0:00
3	142	20.05.2005 0:00
6	142	24.05.2005 0:00
7	142	25.05.2005 0:00
1	142	30.05.2005 0:00
14	142	09.06.2005 0:00
7	142	10.06.2005 0:00
3	142	21.06.2005 0:00
2	142	22.06.2005 0:00
3	142	23.06.2005 0:00
7	142	27.06.2005 0:00
1	142	29.06.2005 0:00
5	142	11.07.2005 0:00
10	142	12.07.2005 0:00
11	142	13.07.2005 0:00
1	142	14.07.2005 0:00
3	142	21.07.2005 0:00
2	142	23.07.2005 0:00
3	142	26.07.2005 0:00
6	142	27.07.2005 0:00
3	142	28.07.2005 0:00
1	142	29.07.2005 0:00
14	142	03.08.2005 0:00
2	142	04.08.2005 0:00
1	142	10.08.2005 0:00
6	142	15.08.2005 0:00
1	142	19.08.2005 0:00
11	142	22.08.2005 0:00
3	142	24.08.2005 0:00
3	142	25.08.2005 0:00
1	142	09.09.2005 0:00
3	142	12.09.2005 0:00
2	142	13.09.2005 0:00
2	142	14.09.2005 0:00
4	142	15.09.2005 0:00
1	142	21.09.2005 0:00

2	142	22.09.2005 0:00
2	142	23.09.2005 0:00
2	142	26.09.2005 0:00
1	142	04.10.2005 0:00
2	142	14.10.2005 0:00
1	142	28.10.2005 0:00
16	142	29.10.2005 0:00
1	142	17.11.2005 0:00
2	142	18.11.2005 0:00
1	142	25.11.2005 0:00
1	142	29.11.2005 0:00
1	142	09.12.2005 0:00
4	142	10.12.2005 0:00
2	142	15.12.2005 0:00
1	142	20.12.2005 0:00
6	142	22.12.2005 0:00
3	142	28.12.2005 0:00
1	142	19.01.2006 0:00
2	142	20.01.2006 0:00
13	142	30.01.2006 0:00
2	142	13.02.2006 0:00
10	142	14.02.2006 0:00
3	142	20.02.2006 0:00
1	142	22.02.2006 0:00
11	142	03.03.2006 0:00
8	142	06.03.2006 0:00
3	142	09.03.2006 0:00
1	142	15.03.2006 0:00
1	142	17.03.2006 0:00
1	142	24.03.2006 0:00
1	142	27.03.2006 0:00
7	142	01.04.2006 0:00
1	142	03.04.2006 0:00
2	142	04.04.2006 0:00
5	142	25.04.2006 0:00
1	142	27.04.2006 0:00
1	142	25.05.2006 0:00
7	142	05.06.2006 0:00
5	142	09.06.2006 0:00
9	142	20.06.2006 0:00
3	142	22.06.2006 0:00
8	142	04.07.2006 0:00
2	142	11.07.2006 0:00
8	142	12.07.2006 0:00
2	142	13.07.2006 0:00
6	142	18.07.2006 0:00
2	142	30.08.2006 0:00
7	142	31.08.2006 0:00
2	142	05.09.2006 0:00
2	142	27.09.2006 0:00
1	142	13.10.2006 0:00
1	142	18.10.2006 0:00
7	142	13.11.2006 0:00
1	142	21.11.2006 0:00
2	142	02.12.2006 0:00
2	142	06.12.2006 0:00
3	142	11.12.2006 0:00
8	142	12.12.2006 0:00
7	142	20.12.2006 0:00
2	142	08.02.2007 0:00
11	142	26.02.2007 0:00
1	142	06.03.2007 0:00
4	142	21.03.2007 0:00
1	142	23.03.2007 0:00
2	142	06.04.2007 0:00
1	142	16.04.2007 0:00
1	142	15.05.2007 0:00
1	142	13.07.2007 0:00
2	142	05.09.2007 0:00
4	142	06.09.2007 0:00
1	142	30.01.2008 0:00
5	142	29.05.2008 0:00
1	142	02.06.2008 0:00
1	142	04.06.2008 0:00
1	142	09.07.2008 0:00
22	142	11.08.2008 0:00
2	142	29.09.2008 0:00
6	142	09.10.2008 0:00
4	142	21.11.2008 0:00

1	142	04.12.2008 0:00
2	142	09.01.2009 0:00
5	142	13.01.2009 0:00
2	142	31.01.2009 0:00
1	142	10.02.2009 0:00
1	142	23.03.2010 0:00
2	142	13.05.2010 0:00
1	142	09.07.2010 0:00
2	142	13.10.2010 0:00
3	142	19.11.2010 0:00
1	142	18.02.2011 0:00
1	142	15.04.2011 0:00
1	142	03.06.2011 0:00
3	142	04.07.2011 0:00
1	142	13.09.2012 0:00
3	142	02.11.2012 0:00
6	142	30.01.2013 0:00
1	142	07.06.2013 0:00
20	142	10.06.2013 0:00
1	142	20.06.2013 0:00
9	142	23.12.2013 0:00
6	142	24.12.2013 0:00
4	142	26.12.2013 0:00
2	142	10.01.2014 0:00
12	142	24.01.2014 0:00
1	142	28.01.2014 0:00
7	142	31.01.2014 0:00
4	142	13.02.2014 0:00
1	142	18.02.2014 0:00
5	142	20.02.2014 0:00
1	142	21.02.2014 0:00
1	142	25.02.2014 0:00
1	142	13.05.2014 0:00
3	142	19.08.2014 0:00
6	142	16.09.2014 0:00
14	142	17.09.2014 0:00
4	142	01.10.2014 0:00

Таблица Е.2 – Данные по изделию 220

Значение	Код изделия	Дата, время
2	220	13.01.2000 0:00
3	220	02.02.2000 0:00
2	220	05.04.2000 0:00
1	220	17.04.2000 0:00
2	220	27.04.2000 0:00
2	220	19.07.2000 0:00
2	220	05.12.2000 0:00
24	220	14.02.2001 0:00
2	220	15.02.2001 0:00
2	220	19.04.2001 0:00
12	220	07.05.2001 0:00
6	220	15.05.2001 0:00
4	220	13.06.2001 0:00
2	220	20.06.2001 0:00
16	220	18.10.2001 0:00
4	220	17.12.2001 0:00
1	220	16.04.2002 0:00
2	220	18.11.2002 0:00
12	220	18.03.2003 0:00
4	220	26.03.2003 0:00
7	220	08.04.2003 0:00
20	220	22.05.2003 0:00
2	220	23.06.2003 0:00
1	220	11.07.2003 0:00
14	220	22.10.2003 0:00
17	220	28.10.2003 0:00
1	220	29.10.2003 0:00
4	220	14.11.2003 0:00
4	220	08.12.2003 0:00
1	220	25.12.2003 0:00
8	220	27.01.2004 0:00
2	220	24.02.2004 0:00

3	220	23.03.2004 0:00
4	220	25.03.2004 0:00
2	220	30.04.2004 0:00
1	220	11.08.2004 0:00
4	220	12.10.2004 0:00
2	220	06.12.2004 0:00
1	220	23.12.2004 0:00
2	220	29.12.2004 0:00
2	220	01.02.2005 0:00
2	220	08.02.2005 0:00
4	220	21.02.2005 0:00
3	220	26.02.2005 0:00
1	220	31.03.2005 0:00
2	220	19.04.2005 0:00
2	220	20.04.2005 0:00
4	220	08.06.2005 0:00
4	220	09.06.2005 0:00
2	220	14.10.2005 0:00
1	220	26.10.2005 0:00
1	220	02.11.2005 0:00
2	220	03.11.2005 0:00
2	220	17.01.2006 0:00
1	220	18.02.2006 0:00
2	220	29.03.2006 0:00
2	220	24.04.2006 0:00
2	220	30.06.2006 0:00
4	220	14.07.2006 0:00
2	220	24.07.2006 0:00
1	220	21.12.2006 0:00
36	220	05.09.2007 0:00
1	220	27.06.2008 0:00
3	220	23.07.2008 0:00
5	220	03.07.2012 0:00
9	220	24.08.2012 0:00
4	220	03.10.2012 0:00
24	220	16.11.2012 0:00
3	220	20.11.2012 0:00
10	220	21.11.2012 0:00
12	220	27.11.2012 0:00
16	220	11.12.2012 0:00
25	220	27.12.2012 0:00
24	220	28.12.2012 0:00
7	220	25.01.2013 0:00
22	220	28.01.2013 0:00
1	220	01.02.2013 0:00
2	220	12.02.2013 0:00
6	220	14.02.2013 0:00
5	220	18.02.2013 0:00
54	220	19.02.2013 0:00
5	220	20.02.2013 0:00
3	220	25.02.2013 0:00
4	220	26.02.2013 0:00
1	220	28.02.2013 0:00
26	220	04.03.2013 0:00
1	220	05.03.2013 0:00
28	220	11.03.2013 0:00
37	220	22.03.2013 0:00
35	220	25.03.2013 0:00
100	220	28.03.2013 0:00
4	220	01.04.2013 0:00
1	220	03.04.2013 0:00
17	220	05.04.2013 0:00
6	220	17.04.2013 0:00
5	220	25.04.2013 0:00
10	220	26.04.2013 0:00
30	220	30.04.2013 0:00
4	220	03.05.2013 0:00
4	220	14.05.2013 0:00
2	220	15.05.2013 0:00

42	220	16.05.2013 0:00
1	220	21.05.2013 0:00
1	220	04.06.2013 0:00
17	220	06.06.2013 0:00
43	220	17.07.2013 0:00
12	220	23.07.2013 0:00
12	220	26.07.2013 0:00
1	220	28.08.2013 0:00
1	220	11.09.2013 0:00
1	220	01.10.2013 0:00
6	220	03.10.2013 0:00
1	220	17.01.2014 0:00
8	220	19.03.2014 0:00
3	220	21.03.2014 0:00
1	220	16.05.2014 0:00
4	220	30.06.2014 0:00

Таблица Е.3 – Данные по изделию 275

Значение	Код изделия	Дата, время
3	275	27.05.1996 0:00
5	275	26.09.1996 0:00
3	275	20.11.1996 0:00
11	275	20.03.1997 0:00
24	275	21.04.1997 0:00
3	275	07.07.1997 0:00
6	275	15.07.1997 0:00
3	275	16.07.1997 0:00
4	275	22.07.1997 0:00
1	275	23.07.1997 0:00
6	275	16.10.1997 0:00
3	275	20.10.1997 0:00
2	275	01.12.1997 0:00
1	275	24.02.1998 0:00
5	275	16.03.1998 0:00
1	275	18.03.1998 0:00
13	275	25.03.1998 0:00
29	275	21.07.1998 0:00
15	275	27.10.1998 0:00
2	275	18.03.1999 0:00
2	275	10.12.1999 0:00
1	275	18.01.2000 0:00
10	275	19.01.2000 0:00
2	275	26.01.2000 0:00
1	275	31.01.2000 0:00
2	275	17.02.2000 0:00
2	275	23.02.2000 0:00
5	275	09.03.2000 0:00
2	275	26.04.2000 0:00
1	275	27.04.2000 0:00
1	275	27.06.2000 0:00
3	275	04.07.2000 0:00
1	275	18.08.2000 0:00
1	275	22.08.2000 0:00
3	275	24.08.2000 0:00
2	275	05.09.2000 0:00
1	275	19.09.2000 0:00
13	275	22.09.2000 0:00
2	275	10.10.2000 0:00
1	275	21.11.2000 0:00
1	275	30.11.2000 0:00
4	275	04.12.2000 0:00
2	275	05.12.2000 0:00
1	275	07.12.2000 0:00
5	275	25.12.2000 0:00
1	275	15.02.2001 0:00
2	275	27.02.2001 0:00
1	275	28.02.2001 0:00
2	275	20.03.2001 0:00
3	275	12.04.2001 0:00
1	275	07.05.2001 0:00
15	275	10.05.2001 0:00
2	275	26.06.2001 0:00
6	275	04.07.2001 0:00
2	275	10.07.2001 0:00

4	275	11.07.2001 0:00
1	275	07.08.2001 0:00
1	275	15.08.2001 0:00
1	275	16.08.2001 0:00
1	275	30.08.2001 0:00
1	275	31.08.2001 0:00
1	275	06.09.2001 0:00
12	275	07.09.2001 0:00
2	275	31.10.2001 0:00
2	275	09.11.2001 0:00
2	275	10.12.2001 0:00
2	275	01.02.2002 0:00
8	275	29.03.2002 0:00
1	275	06.05.2002 0:00
2	275	21.06.2002 0:00
1	275	31.07.2002 0:00
12	275	08.08.2002 0:00
2	275	14.08.2002 0:00
1	275	09.09.2002 0:00
1	275	16.09.2002 0:00
1	275	18.09.2002 0:00
1	275	19.09.2002 0:00
4	275	07.10.2002 0:00
2	275	10.10.2002 0:00
2	275	24.10.2002 0:00
2	275	08.11.2002 0:00
4	275	09.12.2002 0:00
9	275	20.01.2003 0:00
4	275	22.01.2003 0:00
1	275	18.02.2003 0:00
1	275	05.03.2003 0:00
1	275	12.03.2003 0:00
2	275	18.03.2003 0:00
6	275	31.03.2003 0:00
3	275	15.04.2003 0:00
1	275	22.05.2003 0:00
5	275	23.05.2003 0:00
2	275	27.05.2003 0:00
1	275	29.05.2003 0:00
1	275	16.06.2003 0:00
6	275	11.07.2003 0:00
1	275	23.07.2003 0:00
1	275	28.07.2003 0:00
4	275	21.08.2003 0:00
1	275	23.09.2003 0:00
1	275	03.11.2003 0:00
4	275	12.11.2003 0:00
1	275	19.11.2003 0:00
1	275	20.11.2003 0:00
3	275	21.11.2003 0:00
1	275	06.01.2004 0:00
1	275	15.01.2004 0:00
4	275	19.01.2004 0:00
1	275	09.02.2004 0:00
1	275	09.03.2004 0:00
1	275	10.03.2004 0:00
6	275	12.03.2004 0:00
1	275	21.04.2004 0:00
2	275	29.04.2004 0:00
4	275	30.04.2004 0:00
1	275	06.05.2004 0:00
1	275	21.05.2004 0:00
1	275	25.05.2004 0:00
1	275	31.05.2004 0:00
1	275	02.07.2004 0:00
1	275	05.07.2004 0:00
2	275	19.07.2004 0:00
2	275	05.08.2004 0:00
26	275	15.09.2004 0:00
3	275	22.09.2004 0:00
3	275	27.09.2004 0:00
3	275	29.09.2004 0:00
1	275	07.10.2004 0:00
1	275	29.10.2004 0:00
1	275	01.11.2004 0:00
1	275	26.11.2004 0:00
14	275	24.12.2004 0:00
7	275	25.12.2004 0:00
1	275	25.01.2005 0:00

10	275	09.02.2005 0:00
2	275	12.02.2005 0:00
2	275	17.02.2005 0:00
23	275	24.02.2005 0:00
2	275	02.03.2005 0:00
2	275	22.03.2005 0:00
1	275	25.03.2005 0:00
3	275	28.03.2005 0:00
1	275	15.04.2005 0:00
1	275	12.05.2005 0:00
10	275	04.07.2005 0:00
2	275	05.07.2005 0:00
7	275	14.07.2005 0:00
12	275	18.07.2005 0:00
1	275	28.07.2005 0:00
13	275	29.07.2005 0:00
2	275	08.08.2005 0:00
1	275	25.08.2005 0:00
2	275	26.09.2005 0:00
1	275	10.10.2005 0:00
1	275	18.10.2005 0:00
6	275	28.11.2005 0:00
1	275	28.12.2005 0:00
1	275	16.01.2006 0:00
2	275	30.08.2006 0:00
1	275	28.10.2008 0:00
4	275	16.07.2014 0:00
1	275	23.07.2014 0:00
3	275	04.08.2014 0:00
7	275	18.08.2014 0:00
6	275	08.09.2014 0:00
3	275	17.09.2014 0:00
4	275	29.09.2014 0:00
1	275	01.10.2014 0:00

Таблица Е.4 – Данные по изделию 14

Значение	Код изделия	Дата, время
1	14	10.01.2000 0:00
1	14	12.01.2000 0:00
6	14	17.01.2000 0:00
2	14	18.01.2000 0:00
2	14	02.02.2000 0:00
1	14	04.04.2000 0:00
1	14	11.05.2000 0:00
1	14	05.07.2000 0:00
2	14	01.11.2000 0:00
1	14	07.02.2001 0:00
2	14	16.03.2001 0:00
1	14	27.06.2001 0:00
3	14	24.07.2001 0:00
1	14	04.02.2002 0:00
1	14	11.04.2002 0:00
1	14	06.06.2002 0:00
1	14	10.06.2002 0:00
4	14	02.07.2002 0:00
1	14	04.07.2002 0:00
1	14	05.09.2002 0:00
1	14	20.09.2002 0:00
1	14	18.11.2002 0:00
1	14	25.11.2002 0:00
1	14	09.04.2003 0:00
1	14	18.06.2003 0:00
1	14	25.03.2004 0:00
1	14	18.03.2005 0:00
1	14	03.05.2005 0:00
1	14	31.05.2005 0:00
1	14	04.07.2005 0:00

4	14	13.03.2006 0:00
1	14	29.05.2006 0:00
1	14	18.08.2006 0:00
8	14	26.03.2010 0:00
9	14	15.04.2010 0:00
2	14	26.05.2010 0:00
17	14	11.06.2010 0:00
2	14	12.12.2012 0:00
76	14	01.04.2013 0:00
11	14	11.04.2013 0:00
26	14	17.04.2013 0:00
2	14	29.04.2013 0:00
9	14	08.05.2013 0:00
23	14	21.05.2013 0:00
7	14	27.05.2013 0:00
3	14	20.09.2013 0:00
1	14	05.12.2013 0:00
1	14	12.12.2013 0:00
1	14	19.02.2014 0:00
2	14	28.02.2014 0:00

Таблица Е.5 – Данные по изделию 54

Значение	Код изделия	Дата, время
1	54	23.03.2004 0:00
1	54	11.07.2005 0:00
1	54	19.07.2005 0:00
1	54	19.10.2005 0:00
4	54	26.10.2005 0:00
3	54	15.11.2005 0:00
1	54	26.01.2006 0:00
1	54	18.02.2006 0:00
1	54	06.12.2007 0:00
1	54	11.04.2008 0:00
9	54	25.08.2008 0:00
36	54	30.09.2008 0:00
1	54	11.11.2008 0:00
72	54	01.12.2008 0:00
1	54	05.02.2009 0:00
2	54	19.02.2009 0:00
1	54	26.02.2009 0:00
5	54	04.03.2009 0:00
2	54	25.03.2009 0:00
1	54	19.05.2009 0:00
1	54	13.07.2009 0:00
2	54	03.12.2009 0:00
32	54	25.02.2010 0:00
52	54	06.04.2010 0:00
1	54	14.09.2010 0:00
7	54	23.09.2011 0:00
20	54	11.05.2012 0:00
31	54	07.03.2013 0:00
52	54	12.04.2013 0:00
3	54	16.05.2013 0:00
8	54	23.05.2013 0:00
17	54	27.05.2013 0:00
23	54	02.07.2013 0:00
3	54	05.08.2013 0:00
1	54	30.05.2014 0:00
1	54	26.06.2014 0:00

Таблица Е.6 – Данные по изделию 180

Значение	Код изделия	Дата, время
1	180	21.01.2000 0:00
2	180	11.05.2000 0:00
1	180	02.11.2000 0:00
1	180	05.11.2000 0:00
1	180	15.02.2002 0:00
1	180	02.07.2002 0:00
1	180	20.09.2002 0:00
6	180	27.12.2002 0:00
1	180	01.09.2003 0:00
1	180	02.08.2004 0:00
1	180	23.12.2004 0:00
1	180	31.03.2005 0:00
1	180	27.10.2006 0:00
1	180	06.11.2006 0:00
6	180	28.08.2008 0:00
1	180	21.11.2008 0:00
17	180	26.11.2008 0:00
18	180	28.11.2008 0:00
1	180	29.11.2008 0:00
2	180	22.12.2008 0:00
18	180	25.02.2011 0:00
2	180	12.05.2011 0:00
82	180	24.05.2011 0:00
15	180	09.06.2011 0:00
2	180	14.06.2011 0:00
6	180	15.06.2011 0:00
206	180	16.06.2011 0:00
137	180	17.06.2011 0:00
2	180	24.06.2011 0:00
1	180	27.06.2011 0:00
7	180	05.07.2011 0:00
17	180	12.07.2011 0:00
18	180	14.09.2011 0:00
3	180	19.09.2011 0:00
2	180	22.09.2011 0:00
1	180	10.10.2011 0:00
3	180	19.10.2011 0:00
5	180	31.10.2011 0:00
26	180	23.11.2011 0:00
2	180	02.12.2011 0:00
28	180	08.12.2011 0:00
14	180	13.12.2011 0:00
11	180	22.12.2011 0:00
18	180	20.01.2012 0:00
3	180	10.02.2012 0:00
1	180	15.02.2012 0:00
1	180	16.02.2012 0:00
1	180	17.02.2012 0:00
60	180	02.03.2012 0:00
21	180	13.03.2012 0:00
19	180	22.03.2012 0:00
4	180	26.03.2012 0:00
2	180	07.04.2012 0:00
13	180	10.04.2012 0:00
4	180	18.04.2012 0:00
16	180	23.04.2012 0:00
1	180	26.04.2012 0:00
4	180	05.05.2012 0:00
1	180	25.05.2012 0:00
1	180	28.05.2012 0:00
2	180	25.06.2012 0:00
7	180	18.07.2012 0:00
2	180	31.07.2012 0:00
1	180	19.09.2012 0:00
2	180	24.09.2012 0:00
1	180	04.10.2012 0:00

7	180	12.10.2012 0:00
17	180	29.10.2012 0:00
4	180	16.11.2012 0:00
1	180	22.11.2012 0:00
1	180	06.12.2012 0:00
1	180	07.12.2012 0:00
1	180	11.12.2012 0:00
9	180	18.01.2013 0:00
2	180	15.05.2013 0:00
16	180	06.09.2013 0:00
1	180	06.06.2014 0:00
2	180	16.10.2014 0:00

Таблица Е.7 – Данные по изделию 41

Значение	Код изделия	Дата, время
1	41	09.03.2005 0:00
1	41	02.11.2005 0:00
3	41	12.08.2008 0:00
3	41	10.12.2009 0:00
3	41	11.12.2009 0:00
1	41	19.12.2009 0:00
2	41	22.12.2009 0:00
4	41	26.12.2009 0:00
3	41	28.12.2009 0:00
5	41	18.01.2010 0:00
3	41	08.07.2010 0:00
1	41	11.10.2010 0:00
1	41	12.10.2010 0:00
4	41	24.12.2010 0:00
4	41	18.03.2011 0:00
1	41	08.06.2011 0:00
1	41	10.06.2011 0:00
2	41	14.06.2011 0:00
3	41	23.06.2011 0:00
1	41	28.06.2011 0:00
1	41	06.07.2012 0:00
2	41	21.08.2014 0:00
2	41	26.08.2014 0:00
4	41	27.08.2014 0:00

ПРИЛОЖЕНИЕ F. ДАННЫЕ МЕТРИК ПРОЕКТА «РАЗРАБОТКА НЕЧЕТКИХ МОДЕЛЕЙ И РЕАЛИЗАЦИЯ НЕЧЕТКИХ МЕТОДОВ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ»

Таблица F.1 – Данные метрик

#	Трекер	Тема	Назначена	Обновлено
1bafeb7	Bug	first commit	Guskov Gleb	Thu Sep 18 23:31:21 2014 +0400
8f6cce7	Feature	add nls	Anton Romanov	Sat Sep 20 20:17:15 2014 +0400
078aa35	Bug	commit	Timina Irina	Thu Sep 25 22:40:50 2014 +0400
044f5ed	Feature	add dsa	Timina Irina	Mon Sep 29 19:10:42 2014 +0400
155376a	Bug	change AIC for lm model to R-package	Guskov Gleb	Thu Oct 2 18:40:10 2014 +0400
55d302d	Feature	add nls model fix dsa sin	Guskov Gleb	Thu Oct 9 19:47:24 2014 +0400
d1baeb7	Feature	other analysis methods	Timina Irina	Fri Oct 10 22:23:24 2014 +0400
b593f2c	Bug	update dsa	Timina Irina	Sat Oct 11 23:52:15 2014 +0400
2edddd3	Feature	add weighted Aic	Guskov Gleb	Thu Oct 16 13:49:20 2014 +0400
b29530d	Bug	added methods of exponential smoothing	Anton Romanov	Sat Oct 18 20:49:19 2014 +0400
010de2a	Feature	add hqic functions	Timina Irina	Sat Oct 18 22:11:14 2014 +0400
2f4ba77	Bug	Fuzzy Adoptive weights feature + stable project need testing	Guskov Gleb	Mon Oct 20 14:56:36 2014 +0400
6147956	Feature	implement forecast 5.6 R package aic version. add aic, aicc and bic functions	Guskov Gleb	Wed Oct 22 18:06:57 2014 +0400
d786792	Bug	commit	Anton Romanov	Thu Oct 23 20:13:44 2014 +0400
27c95f7	Bug	fuzzy and file	Anton Romanov	Thu Oct 23 20:25:05 2014 +0400
67a640d	Bug	full worked weights models n criteries	Guskov Gleb	Thu Oct 23 21:11:43 2014 +0400
8d2b5c8	Bug	move combinations with weights to separate file	Guskov Gleb	Sat Oct 25 23:15:38 2014 +0400
2ec9a66	Feature	add sbic functions	Timina Irina	Sat Oct 25 23:41:23 2014 +0400
a844b295	Bug	test functions hqic, sbic	Anton Romanov	Tue Oct 28 20:22:31 2014 +0400
30d7c57	Feature	update hqic	Timina Irina	Fri Oct 31 21:14:36 2014 +0400
18fdd85	Feature	update sbic	Timina Irina	Thu Nov 6 22:41:15 2014 +0300
19a203a	Feature	Vovks big update	Guskov Gleb	Fri Nov 14 19:18:14 2014 +0300
7c02b4b	Bug	commit	Anton Romanov	Wed Nov 19 19:56:08 2014 +0300
430d573	Bug	commit	Anton Romanov	Wed Nov 26 05:17:55 2014 +0300
e2d162a	Feature	fix infinity numbers check	Guskov Gleb	Fri Nov 28 14:10:55 2014 +0300
ea86d9c	Bug	worked Vovk	Guskov Gleb	Fri Nov 28 20:19:03 2014 +0300
f305c60	Feature	add log n Box-Cox preparation feature	Guskov Gleb	Fri Dec 5 18:29:41 2014 +0300
3d20988	Feature	fix methods of exponential smoothing	Timina Irina	Sat Dec 6 19:23:12 2014 +0300
0deb260	Bug	commit	Anton Romanov	Tue Dec 9 18:12:48 2014 +0300
905aeac	Bug	changes after testing	Anton Romanov	Fri Dec 12 20:44:59 2014 +0300
1b4a469	Feature	fix error in methods	Timina Irina	Mon Dec 15 17:17:09 2014 +0300
494047d	Feature	fix error in createModel	Anton Romanov	Tue Dec 16 21:37:14 2014 +0300
e09eddc	Bug	commit	Timina Irina	Wed Dec 17 22:31:02 2014 +0300
aa76f5b	Feature	update fix in methods	Anton Romanov	Fri Dec 19 22:51:06 2014 +0300
a4db083	Bug	bug fix in fuzzy methods	Anton Romanov	Wed Dec 24 21:38:25 2014 +0300
7895ac0	Bug	params binary search	Guskov Gleb	Sat Dec 27 03:02:13 2014 +0300
9530d44	Feature	add functions eic	Guskov Gleb	Sat Dec 27 04:10:51 2014 +0300
b8f5474	Feature	test before the big merger	Timina Irina	Thu Jan 8 12:20:11 2015 +0300
25b0d93	Bug	crossvalidation for fuzzy methods	Anton Romanov	Fri Jan 9 14:18:17 2015 +0300
8ce4aef	Feature	clear xLearning and xEstimation fix	Anton Romanov	Fri Jan 9 21:49:32 2015 +0300
2ccf3fe	Bug	big final merge	Guskov Gleb	Sat Jan 10 00:23:57 2015 +0300
35480ca	Bug	results and calc process	Guskov Gleb	Sun Jan 11 13:44:42 2015 +0300
bf7dcb9	Bug	season bug target	Guskov Gleb	Mon Jan 12 21:46:59 2015 +0300
fef8d44	Bug	normal distribution ts for parallel, some bugs, VBScript for graph FY	Guskov Gleb	Wed Jan 14 04:11:52 2015 +0300
f910897	Bug	script	Guskov Gleb	Wed Jan 14 04:12:30 2015 +0300

Продолжение таблицы F.1

#	Трекер	Тема	Назначена	Обновлено
35081cd	Feature	0 remove	Anton Romanov	Wed Jan 14 19:47:33 2015 +0300
98939f0	Feature	update fuzzy methods	Timina Irina	Sat Jan 17 13:24:43 2015 +0300
cb346b5	Bug	commit	Timina Irina	Tue Jan 20 23:10:12 2015 +0300
2b2d32e	Bug	commit	Timina Irina	Wed Jan 28 23:10:12 2015 +0300
027ba62	Feature	update fuzzy methods	Anton Romanov	Fri Jan 30 20:21:04 2015 +0300
4c3c936	Feature	update Model	Anton Romanov	Sat Jan 31 22:47:32 2015 +0300
21b2953	Feature	add Model nleic	Anton Romanov	Tue Feb 3 22:28:21 2015 +0300
0d38913	Feature	add Model leic	Timina Irina	Thu Feb 5 20:10:42 2015 +0300
2afbe00	Bug	Model nleic worked	Timina Irina	Sat Feb 7 18:41:10 2015 +0300
acdb180	Bug	bigest commit from challange too many bags was fixed	Guskov Gleb	Mon Feb 9 16:38:16 2015 +0300
98479e4	Bug	Model leic worked	Anton Romanov	Thu Feb 12 20:41:03 2015 +0300
d71b31b	Bug	test results process	Guskov Gleb	Fri Feb 20 20:41:03 2015 +0300
0b29530	Feature	merge	Guskov Gleb	Fri Feb 27 20:41:03 2015 +0300
d477d77	Bug	test results process	Guskov Gleb	Wed Mar 4 20:42:51 2015 +0300
edc6ab3	Feature	add functions Bootstrap IC	Anton Romanov	Tue Mar 10 16:21:42 2015 +0300
4ad38ca	Feature	add Multi-Step Forecast Model Selection	Anton Romanov	Thu Mar 12 18:14:35 2015 +0300
14e3fd4	Bug	module does not work Bootstrap IC	Timina Irina	Wed Mar 18 19:42:51 2015 +0300
a5304c1	Feature	recheck functions Bootstrap IC	Timina Irina	Wed Mar 18 23:21:03 2015 +0300
a075f19	Bug	test Bootstrap IC	Anton Romanov	Thu Mar 26 19:10:43 2015 +0300
03f6392	Feature	update Bootstrap IC	Anton Romanov	Thu Mar 26 22:40:24 2015 +0300
a0b141d	Feature	test before the big merger	Guskov Gleb	Sat Mar 28 10:03:45 2015 +0300
e9f5b2b	Bug	module does not work Bootstrap IC	Guskov Gleb	Thu Apr 2 20:43:14 2015 +0300
0e27128	Bug	fix error in create Model Multi-Step	Guskov Gleb	Thu Apr 2 22:10:43 2015 +0300
d6cb072	Bug	test Model Multi-Step	Guskov Gleb	Thu Apr 2 23:58:24 2015 +0300
3f4f6e2	Bug	test results process	Anton Romanov	Tue Apr 7 19:10:11 2015 +0300
fcffed7	Feature	update Bootstrap IC	Timina Irina	Sat Apr 11 14:21:47 2015 +0300
68c91f6	Feature	test before the big merger	Timina Irina	Sat Apr 11 18:32:54 2015 +0300
090647c	Feature	check the test results	Timina Irina	Wed Apr 15 14:22:03 2015 +0300
e7dca2b	Feature	upgrade other modules	Timina Irina	Fri Apr 17 23:18:52 2015 +0300
504r022	Bug	commit	Anton Romanov	Mon Apr 20 15:54:14 2015 +0300
e605908	Feature	add KLIC	Guskov Gleb	Fri Apr 24 21:45:02 2015 +0300
3e8171a	Feature	recheck Multi-Step	Guskov Gleb	Fri Apr 24 22:58:03 2015 +0300
233fac3	Bug	fix error in create Multi-Step	Guskov Gleb	Mon Apr 27 22:58:03 2015 +0300
3728i98	Bug	test Multi-Step	Guskov Gleb	Tue Apr 28 11:32:41 2015 +0300
643efb6	Feature	update Multi-Step	Anton Romanov	Fri May 8 17:23:54 2015 +0300
8afeb4a	Bug	module does not work Multi-Step	Guskov Gleb	Tue May 12 20:29:40 2015 +0300
12a86a9	Bug	recheck Multi-Step	Guskov Gleb	Tue May 12 23:20:34 2015 +0300
7ccdf6f	Bug	added methods of KLIC	Timina Irina	Wed May 13 02:44:52 2015 +0300
254e000	Bug	test KLIC	Timina Irina	Sat May 16 15:21:33 2015 +0300
eccb81d	Feature	recheck functions KLIC	Anton Romanov	Wed May 20 23:40:36 2015 +0300
306a109	Bug	fix error in create KLIC	Timina Irina	Sat May 23 15:47:23 2015 +0300
6d027ba	Feature	update KLIC	Guskov Gleb	Sat May 23 20:12:52 2015 +0300
624c3c9	Feature	test before the big merger	Guskov Gleb	Sat May 23 22:54:03 2015 +0300
3621b29	Bug	error loading modules	Timina Irina	Mon May 25 14:12:54 2015 +0300
530d389	Feature	recheck modules	Timina Irina	Mon May 25 19:22:14 2015 +0300
13afe93	Feature	merge	Guskov Gleb	Fri May 29 23:56:54 2015 +0300
eb2749e	Bug	test modules	Guskov Gleb	Sat May 30 00:54:12 2015 +0300
5e7d180	Bug	update modules	Guskov Gleb	Sat May 30 02:23:41 2015 +0300
47df4f9	Feature	check the test results	Anton Romanov	Wed Jun 3 18:41:55 2015 +0300
2a8b078	Feature	test before the big merger	Anton Romanov	Fri Jun 5 19:15:32 2015 +0300
b5b2953	Bug	bigest commit	Guskov Gleb	Tue Jun 16 17:44:54 2015 +0300
f8f03b2	Bug	fixed a bug in the module dsa	Timina Irina	Mon Jun 22 20:23:14 2015 +0300
d950667	Bug	update module dsa	Timina Irina	Tue Jun 23 00:38:22 2015 +0300
5f7c65d	Bug	general testing	Anton Romanov	Wed Jun 24 19:10:36 2015 +0300
7da90f6	Bug	merge successfully	Guskov Gleb	Thu Jun 25 23:17:54 2015 +0300
e1f5dc6	Feature	prepare to paper	Guskov Gleb	Fri Jun 26 22:03:52 2015 +0300

ПРИЛОЖЕНИЕ G. РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Таблица G.1 - Результаты

ВР поступления документов в архив (Код изделия)	ВР-предиктор (Код изделия)	Гипотеза	MSE			Тенденция, интенсивность ($\hat{t}_{t+1}^Y$)
			τ_{t+1}	τ_{t+1}^Y	$\hat{t}_{t+1}^Y$	
275	3	Гипотеза 1	0	0	0	Стабильность
		Гипотеза 2	0,04	0,0398	0,0396	Падение среднее
		Гипотеза 3	0,0835	0,0837	0,084	Стабильность
275	14	Гипотеза 1	0	0	0	Стабильность
		Гипотеза 2	0,04	0,0398	0,0393	Падение среднее
		Гипотеза 3	0,0835	0,0837	0,0844	Стабильность
275	180	Гипотеза 1	0	0	0	Стабильность
		Гипотеза 2	0,04	0,0398	0,0397	Падение сильное
		Гипотеза 3	0,0005	0,0005	0,0005	Стабильность
275	41	Гипотеза 1	0	0	0	Стабильность
		Гипотеза 2	0,04	0,0398	0,0397	Падение сильное
		Гипотеза 3	0,0005	0,0837	0,0839	Стабильность
275	54	Гипотеза 1	0	0	0	Стабильность
		Гипотеза 2	0,04	0,0398	0,0139	Падение сильное
		Гипотеза 3	0,0005	0,0005	0,0109	Стабильность
220	3	Гипотеза 1	0	0	0,0068	Стабильность
		Гипотеза 2	0,0023	0,0023	0,0022	Стабильность
		Гипотеза 3	0	0	0	Стабильность
220	180	Гипотеза 1	0,0003	0,0003	0,0003	Стабильность
		Гипотеза 2	0,0023	0,0023	0,0023	Стабильность
		Гипотеза 3	0	0	0	Стабильность

Продолжение таблицы G.1

ВР поступления документов в архив (Код изделия)	ВР-предиктор (Код изделия)	Гипотеза	MSE			Тенденция, интенсивность ($\hat{\tau}_{t+1}^Y$)
			τ_{t+1}	τ_{t+1}^Y	$\hat{\tau}_{t+1}^Y$	
220	41	Гипотеза 1	0,0003	0,0003	0,0003	Стабильность
		Гипотеза 2	0,0023	0,0023	0,0023	Стабильность
		Гипотеза 3	0	0	0	Стабильность
220	142	Гипотеза 1	0,0003	0,0003	0,0004	Стабильность
		Гипотеза 2	0,0023	0,0023	0,0021	Стабильность
		Гипотеза 3	0	0	0	Стабильность
3	180	Гипотеза 1	0	0,0157	0,0439	Стабильность
		Гипотеза 2	0	0,0157	0,0439	Стабильность
		Гипотеза 3	0	0,0157	0,0439	Стабильность
3	41	Гипотеза 1	0	0,0157	0,0247	Стабильность
		Гипотеза 2	0	0,0157	0,0247	Стабильность
		Гипотеза 3	0	0,0157	0,0247	Стабильность
14	142	Гипотеза 1	0	0,0195	0,1424	Стабильность
		Гипотеза 2	0	0,1221	0,1397	Стабильность
		Гипотеза 3	0,996	0,7361	0,3897	Стабильность
54	142	Гипотеза 1	0,907	0,914	0,996	Стабильность
		Гипотеза 2	0	0,0001	0,0593	Стабильность
		Гипотеза 3	0	0,0001	0,0593	Падение сильное