

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 002.073.04,
созданного на базе Федерального исследовательского центра «Информатика и
Управление» Российской академии наук (ФИЦ ИУ РАН) в соответствии с приказом
Минобрнауки России № 747/нк от 22 июня 2016 года, действующим на момент
защиты диссертации, по диссертации Артемова Алексея Валерьевича
на соискание ученой степени кандидата физико-математических наук.

Диссертация «Математические модели временных рядов с трендом в задачах
обнаружения разладки» в виде рукописи по специальности 05.13.18 –
«Математическое моделирование, численные методы и комплексы программ»
выполнена в Федеральном государственном бюджетном учреждении науки «Институт
проблем передачи информации им. А.А. Харкевича Российской академии наук»
(ИППИ РАН).

Диссертация принята к защите 26 октября 2016 г., протокол № 3.

Соискатель Артемов Алексей Валерьевич, гражданин Российской Федерации,
занимает должность аналитика службы анализа больших данных ООО «Яндекс».

В 2012 году соискатель окончил Федеральное государственное бюджетное
образовательное учреждение высшего образования «Московский государственный
университет имени М.В.Ломоносова» по специальности «Физика», в 2015 – очную
аспирантуру Федерального государственного бюджетного образовательного
учреждения высшего образования «Московский государственный университет имени
М.В.Ломоносова» по специальности 05.13.18 – «Математическое моделирование,
численные методы и комплексы программ».

Научный руководитель – кандидат физико-математических наук, Бурнаев
Евгений Владимирович, доцент Центра по научным и инженерным вычислительным
технологиям для задач с большими массивами данных Автономной некоммерческой
образовательной организации высшего образования «Сколковский институт науки и
технологий»

Официальные оппоненты:

1. Стрижов Вадим Викторович, гражданин РФ, доктор физико-математических
наук, научный сотрудник сектора №17 «Интеллектуальный анализ данных»
Федерального исследовательского центра «Информатика и управление» Российской
академии наук.

Отзыв оппонента положительный, имеются замечания:

1) Квазипериодическими в литературе называют сигналы с изменяющейся
частотой, в то время как в диссертационной работе квазипериодическими называются
временные ряды с изменяющимся профилем цикла.

2) Отсутствуют исследование сходимости алгоритмов оценивания трендов,
исследование несмещенности полученных оценок.

3) Для оценки погрешности тренда в вычислительных экспериментах использован
единственный критерий (средняя относительная среднеквадратичная погрешность), не

приведены среднее абсолютное отклонение, среднеквадратичная ошибка, средняя абсолютная ошибка в процентах.

4) Выбор структуры диссертационной работы, при котором раздел, описывающий математические модели временных рядов с трендом, следует за разделом, содержащим описание ансамблей, представляется нелогичным.

2. Житлухин Михаил Валентинович, гражданин РФ, кандидат физико-математических наук, научный сотрудник отдела теории вероятностей и математической статистики Федерального государственного бюджетного учреждения науки Математический институт им. В. А. Стеклова Российской академии наук.

Отзыв оппонента положительный, имеются замечания:

1) Термин «система с интенсивным программным обеспечением» является не очень удачной калькой с английского. Предлагается заменить его на термин «программные системы».

2) Отмечу, что результат главы 1 полностью следует из формул М. Л. Клепцовой, получившей теорему Гирсанова для фрактального броуновского движения. Таким образом, доказательство оптимального фильтра в главе 1 является в значительной мере техническим.

3) Отсутствует теоретическое исследование предложенного в главе 2 функционала эффективности обнаружения разладки.

4) В главе 2 процедуры обнаружения разладки сравниваются согласно критерию «площадь под кривой точность—полнота», не приводятся результаты их сравнения согласно критериям «средняя задержка в обнаружении разладки при заданном среднем времени до ложной тревоги», «средняя задержка в обнаружении разладки при заданной вероятности ложной тревоги».

5) Хотя для метода фильтрации на основе ядерной регрессии приведен листинг алгоритма, для метода фильтрации на основе фильтра для fVn листинг алгоритма отсутствует.

6) Отсутствует сравнительный анализ точности оценивания тренда при использовании различных ядер в методе на основе ядерной регрессии.

7) Диссертация содержит следующие опечатки:

- Во второй главе отмечается, что во всех формулах принято дискретное время, однако используется обозначение $t \geq 0$, характерное для непрерывного времени.
- В формулах 2.3, 2.4, 2.5 у символов вероятности и математического ожидания пропущены нижние индексы, в связи с чем непонятно, к какому распределению они относятся.
- В формуле 2.11 величину X_t необходимо заменить на ξ_t .

Ведущая организация – Федеральное государственное бюджетное учреждение науки Институт проблем управления им. В. А. Трапезникова Российской академии наук, Москва дала положительное заключение (заключение составил Поляк Борис Теодорович, доктор технических наук, главный научный сотрудник лаборатории №7 ИПУ им. В. А. Трапезникова РАН). Замечания:

1. В работе уделено много внимания случайным процессам с так называемым фрактальным броуновским шумом. Не совсем понятно, почему именно эта модель адекватно описывает реальные сложные временные ряды.
2. В обзоре литературы не всегда полно представлена литература на русском языке. Так, отсутствует ссылка на классическую книгу по разладке И.Никифорова. Более того, некоторые публикации из русских журналов представлены лишь их переводными вариантами (например, известная статья Надарая).
3. Описание результатов вычислительных экспериментов для сравнения различных алгоритмов обнаружения разладки является скудным: не содержит не только комментариев относительно того, как эти результаты соотносятся с известными в литературе, но даже и ссылок на какую-либо подобную литературу.

На диссертацию и автореферат поступили отзывы:

1. МГУ им. М.В.Ломоносова, факультет вычислительной математики и кибернетики, кафедра математических методов прогнозирования, подписал к.ф.-м.н., математик 1-й категории Китов Виктор Владимирович. Замечания:
 - Обозначение момента остановки символом ρ является нестандартным.
 - Используются излишне длинные подстрочные индексы («BAYES»).
 - Отсутствуют какие-либо комментарии по поводу выбора сглаживающих ядер $K_h(\varphi, \psi)$.
2. Федеральное государственное бюджетное образовательное учреждение высшего образования «Национальный исследовательский университет «Высшая школа экономики», факультет компьютерных наук, подписал к.т.н., доцент Игнатов Дмитрий Игоревич. Замечания:
 - Термин «квазипериодический сигнал», используемый в автореферате, является неудачным, более корректным является использование термина «сигнал с квазипериодическим трендом».
 - Отсутствуют примеры процедур обнаружения разладки, которые в тексте автореферата названы «слабыми» детекторами.
3. МГУ им. М.В.Ломоносова, механико-математический факультет, кафедра теории вероятностей, подписал д.ф.-м.н., доцент Шабанов Дмитрий Александрович. Замечаний по существу работы нет.

В дискуссии приняли участие:

- 1) Бернштейн А. В., д.ф.-м.н., профессор, г.н.с. лаб. 4-5 ФИЦ ИУ РАН.
- 2) Дудников Е. Е., д.т.н., профессор, в.н.с. лаб.41 ИПУ РАН.

Соискатель имеет 12 опубликованных работ, из них по теме диссертации опубликовано 7 научных работ общим объёмом 2,5 печатных листа, в том числе 1 статья в научных журналах и изданиях, которые включены в перечень российских рецензируемых научных журналов и изданий для опубликования основных научных

результатов диссертаций (включая 3 работы в зарубежных научных изданиях, входящих в индексы цитирования Web of Science и Scopus).

Соискателем опубликовано 3 работы в материалах всероссийских и международных конференций и симпозиумов.

Наиболее значимые научные работы по теме диссертации:

1. Artemov A. V. Effective signal extraction via local polynomial approximation under long-range dependency conditions // Lobachevskii Journal of Mathematics. — 2016. — Vol. 37. — Issue 1. — Pp. 82–99.
2. Артёмов А. В., Бурнаев Е. В. Оптимальное оценивание сигнала, наблюдаемого во фрактальном гауссовском шуме // Теория вероятностей и ее применения. — 2015. — Т. 60. — No. 1. — С. 163-171.
3. Artemov A. V., Burnaev E. V. Ensembles of Detectors for Online Detection of Transient Changes // Proceedings of The 8th International Conference on Machine Vision (ICMV 2015). — 2015. — P. 98751Z–98751Z-5.
4. Artemov A. V., Burnaev E. V. Nonparametric Decomposition of Quasi- periodic Time Series for Change-point Detection // Proceedings of The 8th International Conference on Machine Vision (ICMV 2015). — 2015. — P. 987520–987520-5.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований:

- разработаны новые математические методы оценки параметров тренда временного ряда по данным измерений, выполненным во фрактальном шуме; разработан новый математический метод обнаружения разладки временного ряда на основе ансамблей известных «слабых» детекторов;
- предложены новые математические модели временных рядов с неизвестными статистическими характеристиками, в том числе:
 - математическая модель временного ряда с квазипериодическим трендом, включающая описание амплитуды, профиля цикла и дисперсии временного ряда;
 - математическая модель временного ряда с трендом и «длинной памятью», включающая локальное описание тренда на основе его разложения по полиномам и описание корреляционной структуры наблюдений, характерное для веб-трафика и ряда других типов данных;
- разработаны новые вычислительные схемы, алгоритмы и численные методы: оценивания параметров тренда временного ряда на основе оптимального фильтра для наблюдений с «длинной памятью»; оценивания параметров временного ряда с квазипериодическим трендом на основе непараметрической регрессии;
- исследованы: эффективность полученных теоретических оценок параметров тренда временного ряда, наблюдаемого во фрактальном гауссовском шуме; эффективность использования предложенных математических моделей временных рядов с трендом (в том числе, квазипериодическим) в задачах оценки

нагрузки программной системы и обнаружения ее разладок и аномалий; эффективность использования процедур обнаружения разладок и аномалий на основе ансамблей «слабых» детекторов;

- предложена интерпретация модели краткосрочной разладки как основной в задачах обнаружения разладки случайного процесса; предложен новый функционал эффективности процедур обнаружения разладки, позволяющий учитывать как ситуацию «ложной тревоги», так и ситуацию «пропуска цели», типичную для задач обнаружения краткосрочных разладок.

Теоретическая значимость исследования обоснована тем, что:

- доказаны теорема 1 о структуре оценки максимального правдоподобия параметров тренда временного ряда, наблюдаемого во фрактальном гауссовом шуме, уточняющая известные в литературе результаты; теорема 2 о структуре Байесовской оценки параметров тренда временного ряда, наблюдаемого во фрактальном гауссовом шуме, для гауссовского и равномерного априорных распределений;
- приведен исчерпывающий аналитический обзор, характеризующий современную проблематику, тенденции и достижения в области математических методов и моделей, направленных на обнаружение отказов программных систем; элементы теории фильтрации для фрактальных динамических систем; элементы теории скорейшего обнаружения разладки, включая известные модели и процедуры;
- выявлены ограничения классического подхода теории скорейшего обнаружения разладки, которая не рассматривает случаи краткосрочных отклонений наблюдаемого временного ряда от невозмущенного состояния, характеризующиеся отсутствием гауссовости, отсутствием информации о параметрах распределений наблюдений до или после разладки, коррелированностью наблюдений, трендами или циклами;
- исследованы теоретические свойства оптимальных оценок параметров тренда случайного процесса, наблюдаемого в шуме с «длинной памятью»; эмпирически изучены эффективность алгоритма оценивания параметров сигнала на основе этих оценок и алгоритма оценивания параметров сигнала на основе непараметрической регрессии; проведено эмпирическое сравнение эффективности алгоритма обнаружения разладок на основе ансамблей «слабых» детекторов в широком ряде модельных условий;
- проведено усовершенствование математической модели схемы наблюдений сигнала в шуме с «длинной памятью»; критерия эффективности процедур обнаружения краткосрочной разладки временного ряда.
- применительно к проблематике диссертации результативно использован математический аппарат стохастического анализа, теории непараметрического оценивания сигналов, методы численной оптимизации выпуклых функций; подходы объектно-ориентированного программирования.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что:

- разработаны и внедрены в эксплуатацию программные модули и компоненты, используемые на производстве для обнаружения отказов сервисов компании «Яндекс» (включая сервисы Яндекс.Музыка, Яндекс.Карты и др.), что подтверждено справкой о внедрении результатов диссертационного исследования;
- численно оценены основные характеристики разработанных методов и алгоритмов, позволяющие спрогнозировать эффект от использования разработанных подходов в практике;
- создана технологическая схема практического применения результатов диссертационного исследования в задачах непрерывного мониторинга больших программных систем;
- представлены детальные рекомендации по эффективному развертыванию и масштабированию разработанного комплекса программ для мониторинга крупных систем с высоким числом наблюдаемых характеристик; отработаны ситуации «холодного старта», одновременной обработки множества характеристик, малых обучающих выборок и т.п.

Оценка достоверности результатов исследования выявила:

- достоверность теоретических результатов гарантируется наличием строгих доказательств, основанных на корректном применении хорошо изученных методов стохастического анализа; совпадением полученных оценок с известными результатами в частных случаях линейных задач;
- экспериментальные результаты снабжены описаниями проведенных экспериментов, допускающими их воспроизводимость; продемонстрировано успешное применение результатов исследования в реальных задачах обнаружения отказов программных систем, что подтверждено справкой о внедрении результатов диссертационного исследования;
- использованы публично доступные выборки данных и апробированные алгоритмы моделирования искусственных наборов данных;
- установлено качественное соответствие оперативных характеристик процедур обнаружения разладки, полученных автором, и представленных в независимых источниках;
- апробированы современные методики измерения нагрузки программных систем.

Личный вклад соискателя состоит в том, что он:

- 1) получил оптимальные оценки параметров тренда временного ряда, наблюдаемого в шуме с «длинной памятью»;
- 2) предложил и исследовал математические модели временных рядов, основанные на применении оптимальных фильтров для наблюдений с «длинной памятью»;

- 3) разработал численные методы фильтрации тренда временного ряда, наблюдаемого в шуме с «длинной памятью»;
- 4) предложил математические модели временных рядов с квазипериодическим трендом;
- 5) разработал алгоритмы и численные методы оценивания параметров временных рядов с квазипериодическим трендом;
- 6) предложил критерий качества процедур обнаружения разладки и численный метод выбора параметров ансамблей «слабых» детекторов для оптимизации этого критерия;
- 7) реализовал алгоритмы обнаружения разладки на основе ансамблей, алгоритмы оценивания параметров временных рядов с трендом (в том числе, квазипериодическим) в составе программного комплекса;
- 8) провел вычислительные эксперименты для оценки эффективности обнаружения разладки с использованием ансамблей;
- 9) подготовил публикации по выполненной работе.

Диссертация охватывает все основные вопросы исследуемой научной задачи и соответствует критерию внутреннего единства, что подтверждается наличием основной идеи, рассмотренной на уровнях теории, алгоритмики и программной реализации; взаимосвязи выводов отдельных частей исследования.

Диссертационный совет пришёл к выводу о том, что диссертация представляет собой научно-квалификационную работу, которая соответствует критериям, установленным Положением о присуждении ученых степеней, утвержденным постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, и принял решение присудить Артемову Алексею Валерьевичу ученую степень кандидата физико-математических наук.

При проведении тайного голосования в заседании участвовало 29 человек (включая 11 присутствовавших докторов наук по специальности защищаемой диссертации), из 42 человек, входящих в состав диссертационного совета. Проголосовали: за присуждение учёной степени - 29, против присуждения учёной степени - 0, недействительных бюллетеней - 0.

Председатель диссертационного совета

Академик РАН

Попков Юрий Соломонович.

Ученый секретарь диссертационного совета

Д.т.н., профессор

Крутько Вячеслав Николаевич



24.02.2014