

ОТЗЫВ

официального оппонента на диссертационную работу Артемова Алексея Валерьевича «Математические модели временных рядов с трендом в задачах обнаружения разладки», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.18 – математическое моделирование, численные методы и комплексы программ.

1. Актуальность темы диссертационной работы

Алгоритмы обнаружения моментов изменения статистических свойств процессов, наблюдаемых в режиме реального времени (обнаружения *разладок*), рассматриваются в литературе с 1960-х гг. и успешно применяются в задачах контроля выпускаемой продукции, обеспечения безопасности сложных технических, информационных и инженерных систем, при мониторинге в финансах и т.д. Хорошо известны и основные принципы построения автоматизированных систем, реализующих алгоритмы обнаружения разладок (аномалий, сбоев и т.п.) с привлечением программных средств. В то же самое время, на практике ввиду особенностей реальных данных естественным образом возникают трудности, связанные с необходимостью спецификации модели разладки. Решению этой задачи для нескольких частных случаев, а также разработке эффективной методологии, инкорпорирующей алгоритмы оценивания параметров наблюдаемых временных рядов, посвящена диссертационная работа Артемова А. В.

Областью применения разработанных в диссертационном исследовании подходов являются сложные современные информационные системы, такие как цифровые системы широкополосной связи, сети передачи данных, интернет-сервисы, информатизированные центры обслуживания, интеллектуальные транспортные системы и т.д. В силу широкого распространения подобных систем задача обнаружения их разладок (отказов, сбоев и т.п.), составляющая тему диссертационного исследования, является, безусловно, актуальной.

2. Содержание работы

Текст диссертации состоит из введения, пяти глав, выводов, заключения и списка литературы, содержащего 108 наименований.

Введение содержит обоснование актуальности проблемы обнаружения отказов программных систем; обзор существующих подходов к решению этой задачи и обоснование затруднений при их использовании в практике; отдельно отмечаются особенности данных, возникающие в практических задачах: наличие сезонных колебаний, длинной памяти, априори произвольных отказов, высокий объем данных.

Первая глава посвящена решению задачи оценивания параметров сигнала по результатам его наблюдений в шуме с длинной памятью. В п. 1.1 дается обзор литературы по моделям длинной памяти и фильтрации сигналов с длинной памятью. П. 1.2 содержит описание модели измерения и постановку задачи оценивания параметра модели; в пп. 1.3—1.5 выводятся аналитические выражения для оценки максимального правдоподобия, Байесовской оценки (в двух различных случаях); в качестве следствия показывается, что момент остановки измерений в Байесовской задаче является неслучайной величиной.

Вторая глава содержит описание алгоритмов обнаружения разладки временного ряда, использующего понятие ансамбля процедур обнаружения разладки, т.е. некоторого алгоритма, подразумевающего их совместное использование. В п. 2.2 приводится обзор основных постановок задач обнаружения разладки и статистик в этих задачах. В п. 2.3. дается обоснование использования ансамблей, их определение, примеры ансамблей. В п. 2.4. приводится обзор критериев эффективности процедур обнаружения разладки и вводится новый функционал качества, описывается алгоритм оптимизации значения этого функционала по обучающей выборке. В п. 2.5 приводятся результаты вычислительных экспериментов и сравнительный анализ эффективности ансамблей и известных процедур обнаружения разладки.

В третьей главе содержится описание математических моделей сезонных временных рядов. В п. 3.2 приводится краткий обзор литературы по фильтрации циклических временных рядов. В п. 3.3 излагается алгоритм фильтрации на основе оценок параметров фрактального броуновского движения со сносом. Затем в п. 3.4 описывается модель сезонного сигнала и алгоритм фильтрации на основе ядерной регрессии, модифицированной для сезонных данных. В п. 3.5 приводится схема обнаружения разладки временного ряда при выделении тренда в режиме реального времени. В п. 3.6. излагаются результаты вычислительных экспериментов, исследующие точность выделения тренда, качество обнаружения разладки на его фоне.

Четвертая глава содержит описание разработанного комплекса программ. П. 4.1. обосновывает сделанные при разработке комплекса архитектурные решения; п. 4.2. описывает структуру комплекса; в п. 4.3. приведено описание алгоритма по умолчанию и обсуждается возможность масштабирования программного комплекса.

Пятая глава посвящена представлению результатов применения разработанных подходов в четырех задачах анализа реальных данных. П. 5.2 содержит описание решения задачи прогнозирования значений финансовых показателей, п. 5.3 посвящен описанию решения задачи оценивания нагрузки сетей передачи данных, п. 5.4 содержит описание решения задачи обнаружения отказов поисковой системы «Яндекс», п. 5.5 описывает решение задачи исследования возможности детектирования изменения режима турбулентного течения.

В заключении сформулированы основные результаты диссертации.

3. Научная новизна

Диссертационная работа содержит ряд новых научных результатов, наиболее значимыми из которых являются:

1. Получены оптимальные оценки параметров тренда случайного процесса, управляемого фрактальным броуновским движением, в том числе оценка максимального правдоподобия и Байесовские оценки; получено аналитическое описание момента оптимальной остановки измерения для нормального априорного распределения в Байесовском случае.
2. Предложен новый алгоритм обнаружения разладки временного ряда на основе совместного использования множества статистик, в ходе вычислительных экспериментов убедительно продемонстрирована его эффективность в сравнении с

известными подходами.

3. Предложены новые методы моделирования сезонных временных рядов, адекватные задачам обнаружения разладки, и в ходе вычислительных экспериментов продемонстрирована эффективность разработанных методов в сравнении с известными подходами.
4. Создан масштабируемый комплекс программ, реализующий предложенные методы оценивания параметров временных рядов и обнаружения их разладок.
5. С помощью разработанного программного комплекса решен ряд задач анализа реальных временных рядов.

4. Достоверность и практическая значимость полученных результатов

Теоретические результаты, полученные в диссертационной работе, являются достоверными в силу применения хорошо изученных подходов стохастического анализа и теории случайных процессов, совпадения полученных соотношений с известными в частных случаях линейных задач. Результаты проведенных вычислительных экспериментов являются достоверными в силу подробного описания их схем и могут быть независимо воспроизведены.

Результаты диссертационного исследования могут применяться при создании компонент автоматизированных информационных систем, используемых для выявления разладок; при оценивании параметров временных рядов, отвечающих характеристикам промышленных информационных систем (посещаемость, объем передаваемого трафика и др.).

5. Замечания по диссертации

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 fBm листинг алгоритма отсутствует.
6. Отсутствует сравнительный анализ точности оценивания тренда при использовании различных ядер в методе на основе ядерной регрессии.

Диссертация содержит следующие опечатки:

1. Во второй главе отмечается, что во всех формулах принято дискретное время, однако

- используется обозначение $t \geq 0$, характерное для непрерывного времени.
2. 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.
 3. В формуле 2.11 величину X_t необходимо заменить на ξ_t .

Указанные замечания не влияют на общую положительную характеристику работы.

6. Заключительная оценка

Диссертация Артемова А. В. представляет собой законченную научно-квалификационную работу, в которой предложена и исследована эффективная схема решения задачи о разладке временных рядов с трендом, имеющая важное значение для создания компонент автоматического обнаружения отказов сложных современных информационных систем. Диссертация содержит важные результаты, имеющие научную новизну и обладающие практической ценностью. Автореферат полностью отражает содержание основного текста.

Считаю, что диссертация удовлетворяет требованиям п. 9 «Положения о порядке присуждения ученых степеней» ВАК РФ, утвержденного постановлением Правительства Российской Федерации № 842 от 24 сентября 2013 года, а ее автор Артемов Алексей Валерьевич заслуживает присуждения ему ученой степени кандидата физико-математических наук по специальности 05.13.18 – математическое моделирование, численные методы и комплексы программ.

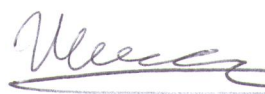
31.01.2017

Официальный оппонент

кандидат физико-математических наук,

научный сотрудник отдела теории вероятностей и математической статистики

Федерального государственного бюджетного учреждения науки «Математический институт им. В. А. Стеклова РАН»

 Житлухин М.В.

Адрес электронной почты: mikhailzh@mi.ras.ru

Телефон: 8 (495) 984-81-41

Федеральное государственное бюджетное учреждение науки Математический институт им. В.А. Стеклова Российской академии наук. Адрес: 119991, Российская Федерация, г. Москва, ул. Губкина, д. 8. Телефон: 8 (495) 984-81-41, <http://www.mi.ras.ru>

