

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА 24.2.437.10,
СОЗДАННОГО НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО
АВТОНОМНОГО ОБРАЗОВАТЕЛЬНОГО УЧРЕЖДЕНИЯ ВЫСШЕГО
ОБРАЗОВАНИЯ «ЮЖНО-УРАЛЬСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ (НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ)» МИНИСТЕРСТВА НАУКИ И ВЫСШЕГО
ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ, ПО ДИССЕРТАЦИИ
НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ КАНДИДАТА НАУК

аттестационное дело № _____

решение диссертационного совета от 22 мая 2024 г. № 1

О присуждении Краевой Яне Александровне, гражданину Российской Федерации, ученой степени кандидата физико-математических наук.

Диссертация «Масштабируемые методы и алгоритмы поиска аномалий во временных рядах» по специальности 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей принята к защите 14 февраля 2024 г. (протокол заседания № 1/п) диссертационным советом 24.2.437.10, созданным на базе федерального государственного автономного образовательного учреждения высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)», Министерство науки и высшего образования Российской Федерации, 454080, г. Челябинск, пр. Ленина, 76, приказ от 12 октября 2022 г. № 1170/нк.

Соискатель Краева Яна Александровна, 10 декабря 1996 года рождения, в 2018 г. и 2020 г. окончила федеральное государственное автономное образовательное учреждение высшего образования «Южно-Уральский государственный университет (национальный исследовательский университет)» по направлениям 02.03.02 «Фундаментальная информатика и информационные технологии» и 02.04.02 «Фундаментальная информатика и информационные технологии». С 01.09.2020 по настоящее время обучается в

очной аспирантуре ФГАОУ ВО «ЮУрГУ (НИУ)» по направлению подготовки высшего образования 09.06.01 «Информатика и вычислительная техника».

Соискатель работает начальником отдела интеллектуального анализа данных и виртуализации Лаборатории суперкомпьютерного моделирования и старшим преподавателем кафедры системного программирования в ФГАОУ ВО «ЮУрГУ (НИУ)» Министерства науки и высшего образования Российской Федерации.

Диссертация выполнена на кафедре системного программирования ФГАОУ ВО «ЮУрГУ (НИУ)», Министерство науки и высшего образования Российской Федерации.

Научный руководитель – доктор физико-математических наук, доцент Цымблер Михаил Леонидович, заместитель директора Научно-образовательного центра «Искусственный интеллект и квантовые технологии» ФГАОУ ВО «ЮУрГУ (НИУ)».

Официальные оппоненты:

- Баркалов Константин Александрович, доктор технических наук, доцент, заведующий кафедрой математического обеспечения и суперкомпьютерных технологий Института информационных технологий, математики и механики федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Нижегородский государственный университет имени Н.И. Лобачевского», Нижний Новгород;
- Никитенко Дмитрий Александрович, кандидат физико-математических наук, ведущий научный сотрудник лаборатории параллельных информационных технологий Научно-исследовательского вычислительного центра федерального государственного бюджетного образовательного учреждения высшего образования «Московский государственный университет имени М.В. Ломоносова», Москва

дали положительные отзывы на диссертацию.

Ведущая организация – федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО», г. Санкт-Петербург, в своем положительном отзыве, подписанном Бухановским Александром Валерьевичем, доктором технических наук, профессором, директором научно-исследовательского института наукоемких компьютерных технологий и утвержденном Никифоровым В.О., доктором технических наук, профессором, проректором на научной работе, указала, что диссертационная работа Я.А. Краевой является законченной научно-квалификационной работой, в которой разработаны масштабируемые методы и алгоритмы поиска аномалий во временных рядах, имеющие существенное значение для решения задач интеллектуального анализа временных рядов на платформе современных высокопроизводительных вычислительных систем. Работа соответствует требованиям Положения о порядке присуждения ученых степеней, в том числе п. 9, а ее автор заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 2.3.5. Математическое и программное обеспечение вычислительных систем, комплексов и компьютерных сетей.

Соискатель имеет 14 опубликованных работ, в том числе по теме диссертации опубликовано 5 работ, из них в рецензируемых научных изданиях опубликовано 4 работы. Все 4 работы опубликованы в российских научных журналах Перечня ВАК (К-1, К-2). По теме диссертации имеется два свидетельства о государственной регистрации программы для ЭВМ. В диссертацию включены результаты, полученные автором лично, авторский вклад составляет 127 стр. (7,9 п.л.).

Публикации полностью соответствуют теме диссертационного исследования и раскрывают ее основные положения. В диссертации отсутствуют недостоверные сведения об опубликованных соискателем ученой степени работах, в которых изложены основные научные результаты диссертации.

Наиболее значимые научные работы соискателя по теме диссертации:

1. Краева, Ya. A Parallel Discord Discovery Algorithm for a Graphics Processor / Ya. Kraeva, M. Zymbler // Pattern Recognition and Image Analysis. – 2023. – Vol. 33, no. 2. – P. 101–112. DOI: 10.1134/S1054661823020062. (Перечень ВАК К1 – Scopus Q3) (12 с./10 с.)
2. Краева, Я.А. Поиск аномалий в сенсорных данных цифровой индустрии с помощью параллельных вычислений / Я.А. Краева // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2023. – Т. 12, № 2. – С. 47–61. DOI: 10.14529/cmse230202. (Перечень ВАК К1 – RSCI)
3. Краева, Я.А. Обнаружение аномалий временного ряда на основе технологий интеллектуального анализа данных и нейронных сетей / Я.А. Краева // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. – 2023. – Т. 12, № 3. – С. 50–71. DOI: 10.14529/cmse230304. (Перечень ВАК К1 – RSCI)
4. Краева, Я.А. Поиск аномалий в больших временных рядах на кластере с GPU узлами / Я.А. Краева, М.Л. Цымблер // Вычислительные методы и программирование. – 2023. – Т. 24, № 3. – С. 291–304. DOI: 10.26089/NumMet.v24r320. (Перечень ВАК К1) (14 с./10 с.)

Свидетельства о регистрации программы

5. Краева, Я.А. DiSSiD: детектор аномалий временного ряда в режиме реального времени / Я.А. Краева // Свидетельство Роспатента о государственной регистрации программы для ЭВМ № 2023685034 от 22.11.2023.
6. Краева, Я.А. PALMAD: детектор аномалий различной длины во временном ряде на графическом процессоре / Я.А. Краева, М.Л. Цымблер // Свидетельство Роспатента о государственной регистрации программы для ЭВМ № 2022667716 от 23.09.2022.

На диссертацию и автореферат поступили положительные отзывы от нижеперечисленных ученых.

1. Газизов Р.К., доктор физ.-мат. наук, профессор, главный инженер проекта, общество с ограниченной ответственностью «РН-БашНИПИнефть», г. Уфа. Замечаний в отзыве нет.
2. Дордопуло А.И., доктор техн. наук, начальник отдела математического и алгоритмического программного обеспечения, общество с ограниченной ответственностью «НИЦ супер-ЭВМ и нейрокомпьютеров», г. Таганрог. Замечания: 1) в автореферате не представлено формальное определение понятия «сниппет»; 2) сравнение разработанных алгоритмов PD3 и PALMAD по времени работы проводилось с алгоритмами Жу и KBF_GPU, а не с алгоритмами DRAG и Merlin, на основе которых они были разработаны.
3. Розенберг В.Л., кандидат физ.-мат. наук, доцент, старший научный сотрудник, федеральное государственное бюджетное учреждение науки Институт математики и механики имени Н.Н. Красовского УрО РАН, г. Екатеринбург. Замечаний в отзыве нет.
4. Ступников С.А., кандидат техн. наук, руководитель отдела, ведущий научный сотрудник, федеральный исследовательский центр «Информатика и управление» РАН, г. Москва. Замечания: 1) следовало бы явно выделить необходимые свойства высокопараллельных вычислительных архитектур, на реализацию в которых нацелены алгоритмы. При замене термина «графический процессор» на «высокопараллельные архитектуры» результаты приобрели бы большую общность; 2) автореферат не включает явные пояснения, почему алгоритмы находят все диссонансы; 3) указание степени загрузки процессоров при экспериментах повысило бы оценку перспектив практической применимости методов и алгоритмов; 4) в автореферате не приведены рекомендации по выбору между аналитическими и нейросетевыми методами для решения конкретных задач поиска аномалий.
5. Черных И.Г., кандидат физ.-мат. наук, ведущий научный сотрудник, федеральное государственное бюджетное учреждение науки Институт

вычислительной математики и математической геофизики СО РАН, г. Новосибирск. Замечаний в отзыве нет.

Выбор официальных оппонентов и ведущей организации обосновывается широкой известностью их научных достижений в областях высокопроизводительных вычислений, параллельного программирования и искусственного интеллекта, направленных на решение различных задач, связанных с анализом временных рядов (мониторинг и выявление нештатных ситуаций в суперкомпьютерных комплексах и др.). В ведущей организации действует Национальный центр когнитивных разработок и исследовательский центр в сфере искусственного интеллекта «Сильный искусственный интеллект в промышленности».

Диссертационный совет отмечает, что **на основании выполненных соискателем исследований:**

- **разработан** комплекс новых параллельных алгоритмов поиска аномалий временного ряда для современных высокопроизводительных кластеров с графическими процессорами, позволяющих существенно сократить время поиска и опережающих известные аналоги по производительности до двух порядков;
- **разработан** новый нейросетевой метод поиска аномалий в потоковом временном ряде, опережающий известные аналоги по точности в среднем на 10 %;
- **разработан и введен** алгоритм автоматизированного формирования обучающей выборки для нейросетевого метода;
- **предложены** оригинальные схемы организации данных и вычислений на основе концепций параллелизма по данным и векторизации вычислений, обеспечивающие существенное увеличение быстродействия алгоритмов поиска аномалий во временном ряде;
- **сформулировано и доказано** утверждение о рекуррентных формулах, которые позволяют сократить объем избыточных вычислений при поиске аномалий во временном ряде.

Теоретическая значимость исследования обоснована тем, что:

- **раскрыта** проблема разработки эффективных методов и алгоритмов для поиска аномалий в больших временных рядах;
- **применительно к проблематике диссертации результативно использованы** в реализации параллельных алгоритмов методы параллельного программирования для графических процессоров и распределенной памяти на основе стандартов CUDA и MPI соответственно, а также методы машинного обучения и интеллектуального анализа данных, теория временных рядов;
- **изучены** существующие современные методы и алгоритмы поиска аномалий во временных рядах;
- **изложены** подходы к обнаружению аномалий во временных рядах: представлены классификация и обзор основных их представителей;
- **теоретически обоснована корректность** предложенных рекуррентных формул, которые позволяют сократить объем избыточных вычислений при поиске аномалий во временном ряде;
- **разработан** новый метод с использованием искусственной нейронной сети для поиска аномалий в потоковом временном ряде и **предложена** новая функция контрастных потерь для ее обучения, обеспечивающие выявление аномалий, отражающих нетипичную и редко встречающуюся активности исследуемого субъекта;
- **проведена модернизация** алгоритма поиска снippetов (типичных подпоследовательностей) временного ряда, повышающая точность поиска и **предложена** эвристика по подбору гиперпараметров для модифицированного алгоритма поиска снippetов.

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

- **созданы** новые масштабируемые методы и алгоритмы поиска аномалий во временных рядах, которые являются агностическими (не имеющими привязки к конкретной предметной области) и могут быть применены в

широком спектре практических приложений при минимальном участии аналитика-эксперта;

- **разработаны и опубликованы в сети Интернет в репозитории GitHub для свободного использования** исходные тексты программ, реализующие предложенные в диссертации разработки;
- **определена эффективность разработанных алгоритмов и методов** в сравнении с известными аналогами путем проведения вычислительных экспериментов на реальных и синтетических временных рядах в соответствии с общепринятыми стандартами;
- **представлены результаты тематических исследований по поиску аномалий в сенсорных данных для предотвращения отказов сложного технологического оборудования в реальных приложениях цифровой индустрии, подтвердившие практическую ценность выполненных разработок.**

Оценка достоверности результатов исследования выявила, что:

- **идея базируется на критическом анализе научных публикаций по теме исследования в высокорейтинговых отечественных и зарубежных журналах;**
- **теория построена на известных методах машинного обучения и интеллектуального анализа временных рядов и методах параллельного программирования для графических процессоров и распределенной памяти;**
- **результаты экспериментов получены на платформе современных высокопроизводительных вычислительных систем: графических процессоров и кластеров с узлами на их основе;**
- **эффективность разработанных алгоритмов и методов подтверждена результатами вычислительных экспериментов, проведенных на синтетических и реальных данных, являющихся де-фактом стандартом для оценки алгоритмов и методов поиска аномалий;**

- **использованы** современные методики проведения экспериментов и анализа их результатов, гарантирующие воспроизводимость экспериментальных результатов;
- **установлено** количественное и качественное соответствие результатов проведенных вычислительных экспериментов исходным данным на основе общепринятых метрик оценки точности обнаружения аномалий во временном ряде.

Личный вклад соискателя состоит в том, что:

- **соискателем** **единолично разработаны** комплекс параллельных алгоритмов поиска диссонансов для современных высокопроизводительных кластеров с графическими процессорами на языке C++ с использованием технологий параллельного программирования MPI и CUDA и нейросетевой метод поиска аномалий в потоковом временном ряде на языке Python.
- в публикациях, выполненных в соавторстве с научным руководителем, научному руководителю принадлежит постановка задачи, **соискателю принадлежат все полученные результаты.**

В ходе защиты диссертации были высказаны следующие **критические замечания:**

1. Традиционный анализ временных рядов основан не только на рассмотрении исходной временной области, но и на привлечении различных частотных характеристик, например, получаемых в результате преобразования Фурье. Поясните, использовались ли в разработанных параллельных алгоритмах частотные методы.
2. В вычислительных экспериментах предложенный соискателем метод DiSSiD сравнивался с известными аналогами, в том числе и с нейросетевыми. Однако в тексте диссертации отсутствует сравнительный анализ результатов с конкурентами, использующими одну из наиболее эффективных архитектур искусственных нейронных сетей, трансформер.

Соискатель Краева Я.А. ответила на задаваемые ей в ходе заседания вопросы и привела собственную аргументацию:

1. В моем исследовании временные ряды не разбивались на частотные компоненты, поскольку разработанные параллельные алгоритмы позволяют находить аномалии в исходных временных рядах.
2. Выбор аналогов для сравнения метода DiSSiD основан на анализе обзорных статей в высокорейтинговых научных журналах, рассматривающих методы и алгоритмы поиска аномалий во временных рядах. Для проведения экспериментов взяты конкуренты, которые в данных обзорах обозначены как обеспечивающие наибольшую точность поиска. Сравнение с трансформерами планируется рассмотреть в будущих исследованиях.

На заседании 22 мая 2024 г. диссертационный совет принял решение за разработку новых масштабируемых методов и алгоритмов поиска аномалий во временных рядах, которые имеют существенное значение для решения задач интеллектуального анализа временных рядов на платформе современных высокопроизводительных вычислительных систем, присудить Краевой Я.А. ученую степень кандидата физико-математических наук.

При проведении тайного голосования диссертационный совет в количестве 11 человек, из них 5 докторов наук по научной специальности диссертации, участвовавших в заседании, из 15 человек, входящих в состав совета, дополнительно введены на разовую защиту 0 человек, проголосовали: за 11, против 0.

Председатель
диссертационного совета



Соколинский Леонид Борисович

Ученый секретарь
диссертационного совета



Иванова Елена Владимировна

22 мая 2024 г.