

ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«ИНФОРМАТИКА И УПРАВЛЕНИЕ»
РОССИЙСКОЙ АКАДЕМИИ НАУК

Гродненский государственный университет
имени Янки Купалы

Отделение математических наук
Российской академии наук

Национальный комитет Российской академии наук
по распознаванию образов и анализу изображений

Интеллектуализация обработки информации

Тезисы докладов 15-й международной конференции
Гродно, 23–27 сентября, 2024 г.

Гродно
ГрГУ им. Янки Купалы
2024

УДК 004.85+004.89+004.93+519.2+519.25+519.7

ББК 22.1:32.973.26-018.2

И 73

Рекомендовано Редакционно-издательским советом
ГрГУ им. Янки Купалы

Редакционная коллегия:

О. В. Сенько, И. А. Матвеев, А. А. Докукин

Интеллектуализация обработки информации: тез. докл. 15-й между-
И 73 нар. конф., Гродно, 23–27 сент. 2024 г. / Федер. исслед. центр «Информатика
и управление» Рос. акад. наук, Гродн. гос. ун-т им. Янки Купалы, Отд. ма-
тем. наук Рос. акад. наук, Нац. ком. Рос. акад. наук по распознаванию образов
и анализу изображений; редкол.: О. В. Сенько, И. А. Матвеев, А. А. Докукин. —
Гродно: ГрГУ, 2024. — 111 с.

ISBN 978-985-582-642-3

В сборнике представлены тезисы докладов 15-й международной конферен-
ции «Интеллектуализация обработки информации», проводимой Федеральным
исследовательским центром «Информатика и управление» Российской акаде-
мии наук и Гродненским государственным университетом имени Янки Купалы,
при организационной поддержке Отделения математических наук Российской
академии наук и Национального комитета Российской академии наук по распо-
знаванию образов и анализу изображений. Конференция проводится регулярно,
начиная с 1989 года, и является представительным научным форумом в области
интеллектуального анализа данных, машинного обучения, распознавания обра-
зов, анализа изображений, обработки сигналов, дискретного анализа. Материал
издания адресован специалистам в области интеллектуализации обработки ин-
формации.

Сайт конференции — <https://mmro.ru>.

УДК 004.85+004.89+004.93+519.2+519.25+519.7

ББК 22.1:32.973.26-018.2

© Авторы докладов, 2024

© Сенько О. В., Матвеев И. А., Докукин А. А., 2024

© ООО «Интеллектуальное кольцо», 2024

© Учреждение образования «Гродненский
государственный университет имени Янки
Купалы», 2024

ISBN 978-985-582-642-3

FEDERAL RESEARCH CENTER
“COMPUTER SCIENCE AND CONTROL”
OF THE RUSSIAN ACADEMY OF SCIENCES

YANKA KUPALA
STATE UNIVERSITY OF GRODNO

DEPARTMENT OF MATHEMATICAL SCIENCES
OF THE RUSSIAN ACADEMY OF SCIENCES

NATIONAL COMMITTEE OF THE RUSSIAN ACADEMY OF SCIENCES
FOR PATTERN RECOGNITION AND IMAGE ANALYSIS

Intelligent Data Processing: Theory and Applications

Book of Abstracts of the 15-th International Conference
Grodno, September 23–27, 2024

Grodno
Yanka Kupala GrSU
2024

UDK 004.85+004.89+004.93+519.2+519.25+519.7
BBK 22.1:32.973.26-018.2

Recommended by the Editorial and Publishing Council
of the Yanka Kupala Grodno State University

Editorial board:
O. V. Senko, I. A. Matveev, A. A. Dokukin

Intelligent Data Processing: Theory and Applications: Book Abstr. 15th Intern. Conf., Grodno, Sept. 23–27, 2024. / Fed. Research Cent. “Inform. and Control” Russian Acad. Sc., Yanka Kupala Grodno St. Uni., Dept. Math. Sc. Russian Acad. Sc., Natl. Comm. Russian Acad. Sc. for Pattern Recognition and Image Analysis; Eds.: O. V. Senko, I. A. Matveev, A. A. Dokukin. — Grodno: GrSU, 2024. — 111 p.

ISBN 978-985-582-642-3

The volume contains the abstracts of the 15th International Conference “Intelligent Data Processing: Theory and Applications”. The conference is held by Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences and Yanka Kupala Grodno State University with organizational support of the Department of Mathematical Sciences of RAS and National Committee of RAS for Pattern Recognition and Image Analysis. The conference has been held regularly since 1989. It is one of the most recognizable scientific forums on data mining, machine learning, pattern recognition, image analysis, signal processing, and discrete analysis. The material of the publication is addressed to specialists in the field of intellectualization of information processing.

The conference website <https://mmro.ru>.

UDK 004.85+004.89+004.93+519.2+519.25+519.7
BBK 22.1:32.973.26-018.2

ISBN 978-985-582-642-3

© Authors of the abstracts, 2024
© Senko O. V., Matveev I. A., Dokukin A. A., 2024
© Intellectual Ring LLC, 2024
© Educational Institution
“Yanka Kupala Grodno State University”, 2024

Оргкомитет

- Председатель:** Матвеев Иван Алексеевич, *д. т. н.*
- Заместители председателя:** Гринь Александр Александрович, *д. ф.-м. н.*
Докукин Александр Александрович, *к. ф.-м. н.*
- Члены комитета:** Егорова Евгения Кирилловна, *к. ф.-м. н.*
Инякин Андрей Сергеевич, *к. ф.-м. н.*
Кадан Александр Михайлович, *к. т. н.*
Карканица Анна Викторовна, *к. т. н.*
Проневич Андрей Францевич, *к. ф.-м. н.*
Рейер Иван Александрович, *к. т. н.*
Борисова Татьяна Игоревна
Громов Андрей Николаевич
Дюкова Анастасия Петровна

Программный комитет

- Председатель:** Соколов Игорь Анатольевич, *акад. РАН*
- Заместители председателя:** Воронцов Константин Вячеславович, *д. ф.-м. н.*
Сенько Олег Валентинович, *д. ф.-м. н.*
- Члены комитета:** Семенов Алексей Львович, *акад. РАН*
Сойфер Виктор Александрович, *акад. РАН*
Шананин Александр Алексеевич, *акад. РАН*
Петров Игорь Борисович, *чл.-корр. РАН*
Хачай Михаил Юрьевич, *чл.-корр. РАН*
Гимади Эдуард Хайрутдинович, *д. ф.-м. н.*
Горнов Александр Юрьевич, *д. т. н.*
Громова Ольга Алексеевна, *д. м. н.*
Двоенко Сергей Данилович, *д. ф.-м. н.*
Дюкова Елена Всеволодовна, *д. ф.-м. н.*
Краснопрошин Виктор Владимирович, *д. т. н.*
Лазарев Александр Алексеевич, *д. ф.-м. н.*
Местецкий Леонид Моисеевич, *д. т. н.*
Пытьев Юрий Петрович, *д. ф.-м. н.*
Стрижов Вадим Викторович, *д. ф.-м. н.*
Чуличков Алексей Иванович, *д. ф.-м. н.*
Ватолин Дмитрий Сергеевич, *к. ф.-м. н.*
Гнеушев Александр Николаевич, *к. ф.-м. н.*
Копылов Андрей Валерьевич, *к. т. н.*
Середин Олег Сергеевич, *к. ф.-м. н.*

Organizing Committee

Chairman: Ivan Matveev, *Dr. Sc. (Tech.)*

Vice-chairman: Alexander Grin, *Dr. Sc. (Phys.-Math.)*
Alexander Dokukin, *Cand. Sc. (Phys.-Math.)*

Committee members: Evgeniya Egorova, *Cand. Sc. (Phys.-Math.)*
Andrei Inyakin, *Cand. Sc. (Phys.-Math.)*
Alexander Kadan, *Cand. Sc. (Tech.)*
Anna Karkanitsa, *Cand. Sc. (Tech.)*
Andrei Pranevich, *Cand. Sc. (Phys.-Math.)*
Ivan Reyer, *Cand. Sc. (Tech.)*
Tatiana Borisova
Andrei Gromov
Anastasija Djukova

Program Committee

Chairman: Igor Sokolov, *Academician of RAS*,

Vice-chairmen: Konstantin Vorontsov, *Dr. Sc. (Phys.-Math.), Prof. RAS*
Oleg Senko, *Dr. Sc. (Phys.-Math.)*

Committee members: Alexey Semenov, *Academician of RAS*
Viktor Soyfer, *Academician of RAS*
Aleksandr Shananin, *Academician of RAS*
Igor Petrov, *corr. member of RAS*
Michael Khachay, *corr. member of RAS*
Edward Gimadi, *Dr. Sc. (Tech.)*
Alexander Gornov, *Dr. Sc. (Tech.)*
Olga Gromova, *Dr. Sc. (Med.)*
Sergei Dvoenko, *Dr. Sc. (Phys.-Math.)*
Elena Djukova, *Dr. Sc. (Phys.-Math.)*
Viktor Krasnoproshin, *Dr. Sc. (Tech.)*
Alexander Lazarev, *Dr. Sc. (Phys.-Math.)*
Leonid Mestetskiy, *Dr. Sc. (Tech.)*
Yury Pytiev, *Dr. Sc. (Tech.)*
Vadim Strizhov, *Dr. Sc. (Phys.-Math.)*
Alexey Chulichkov, *Dr. Sc. (Tech.)*
Dmitriy Vatolin, *Cand. Sc. (Phys.-Math.)*
Aleksandr Gneushev, *Cand. Sc. (Phys.-Math.)*
Andrej Kopylov, *Cand. Sc. (Phys.-Math.)*
Oleg Seredin, *Cand. Sc. (Phys.-Math.)*

Рецензенты

Адуенко А. А.
Анциперов В. Е.
Бахтеев О. Ю.
Визильтер Ю. В.
Володин С. Е.
Воронцов К. В.
Гасников А. В.
Гнеушев А. Н.
Гончаров А. В.
Грабовой А. В.
Двоенко С. Д.
Докукин А. А.
Драгунов Н. А.
Ивахненко А. А.

Инякин А. С.
Исаченко Р. Г.
Катруца А. М.
Копылов А. В.
Қушнир О. А.
Ланге М. М.
Майсурадзе А. И.
Масляков Г. О.
Матвеев И. А.
Местецкий Л. М.
Мурашов Д. М.
Неделько В. М.
Нейчев Р. Г.
Новик В. П.

Одиноких Г. А.
Панов М. Е.
Потапенко А. А.
Рейер И. А.
Сафонов И. В.
Сенько О. В.
Середин О. С.
Стрижов В. В.
Сулимова В. В.
Торшин И. Ю.
Турдаков Д. Ю.
Хачай М. Ю.
Чуличков А. И.
Янина А. О.

Reviewers

Aduenko A.
Antsiperov V.
Bakhteev O.
Chulichkov A.
Dvoenko S.
Dokukin A.
Dragunov N.
Gasnikov A.
Gneushev A.
Goncharov A.
Grabovoy A.
Inyakin A.
Isachenko R.
Katrutsa A.

Ivakhnenko A.
Khachay M.
Kopylov A.
Kushnir O.
Lange M.
Maysuradze A.
Maslyakov G.
Matveev I.
Mestetskiy L.
Murashov D.
Nedel'ko V.
Nejchev R.
Novik V.
Odinokikh G.

Panov M.
Potapenko A.
Reyer I.
Safonov I.
Sen'ko O.
Seredin O.
Strizhov V.
Sulimova V.
Torshin I.
Turdakov D.
Vizil'ter Yu.
Volodin S.
Vorontsov K.
Yanina A.

Краткое оглавление

Интеллектуальный анализ данных	11
Машинное обучение	18
Прогнозирование и рекомендательные системы	31
Аналитика больших данных	42
Нейронные сети и глубокое обучение	44
Методы оптимизации для интеллектуального анализа данных	55
Вычислительная сложность и приближенные методы	59
Обработка и анализ изображений, компьютерное зрение	61
Обработка и анализ сигналов	76
Информационный поиск и анализ текстов	81
Индустриальные приложения науки о данных	87
Анализ биомедицинских данных, биоинформатика	98

Brief Table of Contents

Data Mining	11
Machine Learning	18
Forecasting and Recommendation Systems	31
Big Data Analytics	42
Neural Networks and Deep Learning	44
Data Mining Optimization Techniques	55
Algorithmic Complexity and Approximate Methods	59
Image Processing, Computer Vision	61
Signal Processing	76
Information Retrieval and Text Analysis	81
Industrial Data Science Applications	87
Analysis of Biomedical Data, Bioinformatics	98

Нейроморфное кодирование/декодирование потоков данных на основе модели точечных процессов Пуассона

*Анциперов Вячеслав Евгеньевич*¹*

antciperov@cplire.ru

*Кершнер Владислав Андреевич*¹

vladkershner@mail.ru

¹Москва, Институт радиотехники и электроники им. В. А. Котельникова РАН

Нейроморфные вычисления, основанные на нейробиологических принципах, в настоящее время являются наиболее перспективным направлением для решения сложных информационных проблем, связанных, прежде всего, с огромными объемами накопленных данных. Основываясь на высокопараллельной обработке информации, процедурах, встроенных непосредственно в блоки данных, масштабируемости, событийно-управляемых вычислениях и т. д., новое поколение нейроморфных компьютеров (иногда называемых импульсными (спайковыми) нейронными сетями третьего поколения) потенциально сможет использоваться для хранения чрезвычайно больших объемов цифровых данных и их обработки за приемлемое время и с гораздо меньшим энергопотреблением.

Наряду с ожидаемыми потенциальными характеристиками и многообещающими возможностями нейроморфные вычисления несут в себе и труднопредсказуемые вызовы. Нейроморфные структуры, которые по определению моделируют структуры мозга, будут представлены искусственными нейронами, синапсами и их интенсивно чередующимися связями, в отличие от процессоров и блоков памяти традиционных компьютеров, разделенных общей шиной. Поэтому в нейроморфных системах обработка данных и память должны осуществляться «по месту» нейронной сетью и, соответственно, программы должны задаваться не последовательностями инструкций, а определяться структурой и параметрами нейронной сети. Это открывает новые горизонты, позволяющие не только сосредоточиться на цифровых вычислениях, но и переосмыслить использование аналоговых, приближенных и смешанных вычислений, характерных для биологических нейронов. При этом нейроморфные вычисления потребуют радикального изменения концепции программирования. Учитывая это, в последнее время нами были предприняты попытки разработать некоторые методы работы с потоками данных на основе принципов нейроморфных вычислений [1]. В данной работе представлены некоторые результаты усилий, предпринятых в этом направлении.

Основные черты нашего подхода к нейроморфным вычислениям изложены, например, в [2]. Существенной отправной точкой предлагаемого нами подхода является особое (выборочное) представление входных данных в виде потока дискретных событий (отсчетов), подобных событиям срабатывания нейронов сетчатки. Учитывая специфику выборочного представления, мы сформировали генеративную модель для первичной обработки потока отсчетов в виде семейства параметрических распределений вероятностей, представленных смесями компонентов. Этот выбор мотивирован известными нейрофизиологическими

фактами о системе рецептивных полей (РП) сетчатки (определяемых ганглиозными клетками сетчатки — ГКС), которые реализуют универсальные механизмы (включая латеральное торможение) биологической нейронной сети (мозга). В данной статье обсуждаются новые результаты, касающиеся анализа данных, закодированных на сетчатке, и формирования на их основе того, что Марр назвал *первичным наброском изображения* и который по существу является первым шагом на пути к *Гештальт-интерпретации* входных информационных потоков.

Работа выполнена на средства бюджетного финансирования ИРЭ имени В. А. Котельникова РАН в рамках гос. контракта «Рэлдис-2».

- [1] *Antsipero V. E. Generative Model of Autoencoders Self-Learning on Images Represented by Count Samples // Automation and Remote Control. — 2022. — Vol. 83, No. 12. — Pp. 1959–1983.*
- [2] *Antsipero V. E. Neuromorphic Encoding / Reconstruction of Images Represented by Poisson Counts // Proc. 13th Int. Conf. Pattern Recognition Applications and Methods. — Setubal: SCITEPRESS, 2024. — Pp. 5–25.*

Анализ и интерпретация данных как технология получения знаний о реальности

Чуличков Алексей Иванович¹

achulichkov@gmail.com

¹Москва, МГУ имени М. В. Ломоносова

В докладе рассматриваются математические методы анализа и интерпретации данных, предназначенные для получения наиболее точных оценок параметров изучаемого объекта. Основой рассмотрения являются методы, созданные в научной школе профессора Юрия Петровича Пытьева. В этих методах предполагается, что данные получены в результате измерительного эксперимента, проводимого по схеме $\xi = Af + \nu$, где данные ξ интерпретируются как искаженный шумом ν выходной сигнал измерителя A , на вход которого подан сигнал f от измеряемого объекта. Результатом интерпретации данных ξ является оценка выходного сигнала $u = Uf$ «идеального измерителя» U , на вход которого подан тот же сигнал f от измеряемого объекта, но выходной сигнал есть параметры «ненаблюдаемого» изучаемого объекта. Отметим основные положения, лежащие в основе этих методов, позволяющие утверждать, что результат их применения действительно дает адекватное представление об изучаемом объекте. Во-первых, методы строятся так, чтобы получить максимальную точность оценок параметров $u = Uf$ изучаемого объекта. Во-вторых, поскольку точность оценок определяется математической моделью M , описывающей исследуемый объект, связь его параметров с анализируемыми данными, погрешности данных ν и пр., то в предлагаемом подходе разрабатываются методы, позволяющие проверить согласованность используемой математической модели с анализируемыми

данными, а также согласованность результатов анализа и интерпретации с данными измерений ξ . Для ответа на последний вопрос используется множество альтернативных моделей измерения $\{\widetilde{M}\}$ и вычисляется погрешность $H(M, \widetilde{M})$ оценки Uf , построенной по модели M при условии, что истинной моделью является \widetilde{M} . В качестве гипотезы выдвигается предположение, что выполнено неравенство $H(M, \widetilde{M}) \leq H(M, \widetilde{M})$, означающее, что истинная модель такова, что использование заданной модели M вместо истинной \widetilde{M} не приведет к завышению точности оценки параметров Uf .

К настоящему времени созданы и успешно применяются линейные методы анализа и интерпретации данных [1]. В докладе рассматриваются методы, основанные на нелинейных оценках и применяемые для решения задач дистанционного зондирования, для анализа динамики нелинейных систем, морфологического анализа и интерпретации изображений и др., в том числе методы, использующие технологии искусственных нейронных сетей.

- [1] *Пытьев Ю. П.* Методы математического моделирования измерительно-вычислительных систем. — М.: Изд-во Московского университета, 2024. — 431 с.

Анализ социального настроения городского населения с помощью оценок спектральных плотностей с заданной точностью

Семенчук Наталья Владимировна^{1*}

senata155@gmail.com

*Рудикова-Фронхёфер Лада Владимировна*¹

lada.rudikowa@gmail.com

*Дейцева Анна Геннадьевна*¹

dejceva_ag@grsu.by

¹Гродно, Гродненский государственный университет имени Янки Купалы

Работа посвящена исследованию тенденций и особенностей функционирования данных в виде временных рядов для социального настроения городского населения с помощью оценок спектральных плотностей, построенных по пересекающимся интервалам наблюдений с заданной степенью точности [1]. Объектом исследования являются статистические данные социальных опросов, касающиеся удовлетворенности жизнью городского населения в виде дискретных временных рядов. Предметом исследования являются методы спектрального анализа. Цель данной работы — апробация алгоритмов для выявления периодических компонент (особенностей функционирования) временного ряда социальных данных.

В работе анализировались временные ряды, построенные по данным, собранным с 2013 по 2024 гг. (по месяцам для различных социологических опросов). Для выявления периодичностей в данных применялся метод определения числа интервалов разбиений и количества наблюдений в них при построении оценок

спектральных плотностей стационарных случайных процессов с заданной точностью по пересекающимся интервалам наблюдений на основе асимптотических результатов, полученных для скорости сходимости первого момента в предположении, что спектральная плотность удовлетворяет условию Липшица. Рассматриваются два случая: с единичным и произвольным окном просмотра данных [2].

Данная задача возникла из необходимости структурно анализировать данные в виде стационарных случайных процессов различной длины в автоматизированном режиме, в частности, для спектрального анализа данных социологической направленности и др. Отметим, что графики оценок спектральной плотности помогают раскрыть почти периодические компоненты спектра и их месторасположение, а также могут быть полезны в вопросах выбора модели для дальнейшего прогнозирования поведения исследуемых данных [3].

Исходные временные ряды были обработаны: стандартизированы и приведены к стационарному виду. Далее для каждого ряда были построены оценки спектральных плотностей. Анализируя пики построенных оценок, для каждого временного ряда были определены периодичности, в случае их наличия.

Результат работы алгоритма вполне приемлемый. Следует отметить, что оценивание спектральных плотностей случайных процессов является большим искусством. Применение различных окон просмотра данных, весовых функций, усреднение реализаций в определенных случаях может привести к очень большому выигрышу в точности оценок. Однако, при применении различных методов построения важно помнить о той грани, за которой вместо дальнейшего улучшения оценки наступает ее ухудшение. Полученные результаты могут помочь исследователю в выборе параметров оценивания, а именно числа интервалов разбиения и количества наблюдений в них при построении оценок спектральных плотностей по пересекающимся интервалам наблюдений, основываясь не на интуиции, а на результатах, касающихся скорости сходимости моментов таких оценок. Предложенный метод будет полезен исследователю при анализе данных в виде стационарных случайных процессов с помощью непараметрических методов спектрального анализа в автоматизированном режиме. Достоинством данных методов является их простота, высокая скорость работы и способность адаптироваться под нестандартное поведение данных, что позволяет эффективно выявлять периодичности в динамике исследуемых временных рядов.

- [1] *Андерсон Т. В.* Статистический анализ временных рядов. — М.: Мир, 1976. — 755 с.
- [2] *Труш Н. Н.* Асимптотические методы статистического анализа временных рядов. — Минск: БГУ, 1999. — 218 с.
- [3] *Бриллинджер Д.* Временные ряды. Обработка данных и теория. — М.: Мир, 1980. — 536 с.

Ослабленные обобщенные прецеденты и критерии отбора вторичных кластеров

Ангальт Елена Михайловна¹

angalt2010@mail.ru

Виноградов Александр Петрович^{2*}

vngrccas@mail.ru

¹Оренбург, Оренбургский государственный аграрный университет

²Москва, Федеральный исследовательский центр «Информатика и управление» РАН

Поведение больших выборок управляется объективными закономерностями, связывающими наблюдаемые признаки отдельных объектов, записей, изображений. В. В. Рязановым был предложен подход с обобщенными прецедентами (ОП) в целях компактного представления структуры выборки, когда описание обнаруженного кластера ограниченным набором параметров становится новым объектом [1, 2].

Подход можно рассматривать в качестве многомерного аналога преобразования Хафа. Кластер здесь также представлен точкой в пространстве параметров [3, 4]. В стандартных вариантах преобразования Хафа в 2D и 3D идеалом является пик плотности вторичного распределения, собранный из локальных проявлений искомой формы, но для реальных изображений и сцен нередко возникает сложная вторичная кластерная структура, также заслуживающая исследования.

Ослабленные ОП нацелены на исследование именно вторичного распределения, методы редукции размерности могут варьироваться. Ограниченное число наблюдаемых или производных параметров задает слоение исходного пространства R^N , и в каждом слое отыскиваются локальные корреляции параметров в виде кластеров плотности. Поскольку невозможно задать форму подмножества S в слое для неизвестной пока закономерности, предлагается использовать для селекции вариантов объединения повышенных плотностей в кластеры различные условия «естественности», выполнения которых можно ожидать для S в реальных выборках [5]. Например:

- а) суммарная локальная кривизна S в слое;
- б) суммарная локальная кривизна графика функции эмпирической плотности над S ;
- в) полная вариация (интеграл модуля градиента) эмпирической плотности на S ;
- г) полная вариация градиента плотности на S (и далее для старших производных);
- д) мощность подмножества точек смены знака градиента; и т. д., в зависимости от специфики задачи. Минимизация подобного критерия может служить важным инструментом при выборе оптимального варианта аппроксимации S .

[1] Nelyubina E, Ryzanov V., Vinogradov A. Transforms of Hough Type in Abstract Feature Space: Generalized Precedents // Proc. 12th Int. Joint Conf. Computer Vision,

Imaging and Computer Graphics Theory and Applications VISIGRAPP 2017. — Porto, Portugal: INSTICC, 2017. — Pp. 651–656.

- [2] *Ryazanov V., Vinogradov A.* Analogues of Image Analysis Tools in the Problems of Finding Latent Regularities in Big Applied Data // Pattern Recognition and Image Analysis. — 2022. — Vol. 32, No. 3. — Pp. 639–644.
- [3] *Zhuravlev Yu. I., Ryazanov V. V., Sen'ko O. V., Dokukin A. A., Vinogradov A. P., Nelyubina E. A., Stefanovskii D. V.* Using Hough-Like Transforms for Extracting Relevant Regularities from Big Applied Data // Pattern Recognition and Image Analysis. — 2021. — Vol. 31, No. 4. — Pp. 699–709.
- [4] *Nelyubina E., Ryazanov V., Vinogradov A.* Analogs of Image Analysis Tools in the Search of Latent Regularities in Applied Data // Pattern Recognition, Computer Vision, and Image Processing. ICPR 2022 Int. Workshops and Challenges, Part II, Lecture Notes in Computer Science 13644. — 2023. — Pp. 529–540.
- [5] *Рязанов В. В., Виноградов А. П.* Подходы к выбору содержательных аспектов при анализе сложных выборок // Сборник трудов 21-й Всероссийской конференции с международным участием «Математические методы распознавания образов» (ММО-2023). — М.: Изд-во РАН, 2023. — С. 6–8.

Детекция аномалий в рядах событий при условии высоких требований ко времени обработки

Котов Александр Евгеньевич^{1,2*}

kotov.ae@phystech.edu

*Матвеев Иван Алексеевич*²

matveev@frccsc.ru

¹Москва, МФТИ (НИУ)

²Москва, Федеральный исследовательский центр «Информатика и управление» РАН

Рассматриваются существующие методы и предлагается альтернативное DL-решение задачи детекции аномалий в последовательности событий одной природы и единой идентификации. Особенностью задачи является ограничение на время выдачи ответа. Данная задача актуальна в системах фрод-детекции, антиспам системах, системах защиты В2С-процессов.

Основными методами при таком ограничении являются методы классического машинного обучения: Gradient Boosting, Random Forest, Logistic Regression. Эти методы используют агрегированные признаки ряда и описание последнего события, по которому требуется принять решение.

Предлагаемые подходы основаны на модели глубокого обучения. Для принятия решения методы используют признаки каждого события ряда за определенный период и агрегированные признаки ряда совместно с описанием последнего события.

На первом этапе [1] в разработанном методе используется подход контрастного обучения, целью которого является адаптация весов под домен дан-

ных и детекция неувязок с высокой полнотой, но, возможно, низкой точностью. В качестве базовых архитектур выбраны RRN, transformers и S4.

Вторым этапом выполняется обучение на предсказание аномалии, при этом ряд обрезается с конца настолько, чтобы была возможность регулярного запуска модели при промышленном использовании. Целью отбрасывания последних событий является требование уложиться во время ответа алгоритма. На этом этапе формируется вектор-представление ряда.

На последнем этапе обучается быстрый алгоритм, использующий вектор-представления ряда и оставшиеся признаки (агрегации и описания). Ответ на данном этапе является финальным и отдается в систему с округлением по порогу, исходя из специфики рисков исходной задачи.

Разработка позволила увеличить полноту, при сохранении полноты нулевого класса. Время задержки ответа при описанном подходе не увеличилось.

- [1] *Babaev D., Ovsov N., Kireev I., Ivanova M., Gusev G., Nazarov I., Tuzhilin A.* CoLES: Contrastive Learning for Event Sequences with Self-Supervision // Proc. 2022 Int. Conf. Management of Data, 2022. — Pp. 1190–1199.

О теоретико-информационных границах точности в задачах кодирования источников сообщений и анализа данных

Ланге Михаил Михайлович¹

lange_mm@mail.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» РАН

Исследуются нижние границы средней погрешности как функции количества используемой информации для вероятностных моделей кодирования дискретных и непрерывных сообщений, переданных по каналу с шумом. Приводятся аналогичные границы для моделей классификации данных и оценивания случайного параметра распределения.

Модель кодирования сообщений представлена парой стохастических преобразований $U \rightarrow V \rightarrow \hat{U}$, где U — множество независимых дискретных или непрерывных сообщений с заданным распределением или плотностью распределения; V — множество сообщений на выходе стохастического канала без памяти; \hat{U} — множество оценок сообщений источника с погрешностью $d(u, \hat{u}) \geq 0$, $u \in U$, $\hat{u} \in \hat{U}$. Условные по $v \in V$ распределения оценок $\hat{u} \in \hat{U}$ образуют множество $Q_{\hat{U}|V}$. Вероятностные характеристики модели и мера погрешности позволяют ввести среднюю взаимную информацию $I_{Q_{\hat{U}|V}}(V; \hat{U})$ и среднюю погрешность $E_{Q_{\hat{U}|V}}(V; \hat{U})$ в форме функционалов, зависящих от $Q_{\hat{U}|V}$ [1].

Для дискретного источника, когда $\|U\| = \|\hat{U}\| = q \geq 2$, $\|V\| \geq q$ и меры Хемминга $d(u, \hat{u})$, определяемой индикатором $[u \neq \hat{u}]$, в [1] найден минимум $I_{Q_{\hat{U}|V}}(V; \hat{U})$ по распределениям $Q_{\hat{U}|V}$ при условии $E_{Q_{\hat{U}|V}}(V; \hat{U}) \leq \varepsilon$ в форме монотонно убывающей функции

$$R(\varepsilon) = I(U; V) - h(\varepsilon - \varepsilon_{\min}^{(q)}) - (\varepsilon - \varepsilon_{\min}^{(q)}) \ln(q - 1), \quad \varepsilon_{\min}^{(q)} \leq \varepsilon \leq \varepsilon_{\max}^{(q)}. \quad (1)$$

Здесь $h(z) = -z \ln z - (1 - z) \ln(1 - z)$, $R(\varepsilon_{\min}^{(q)}) = I(U, V)$ и $R(\varepsilon_{\max}^{(q)}) = 0$, причем, $\varepsilon_{\min}^{(q)}$ уменьшается с ростом взаимной информации $I(U, V)$. Для непрерывного источника и меры погрешности $d(u, \hat{u}) = (u - \hat{u})^2$ в [2] получена аналогичная зависимость в форме

$$R(\varepsilon) = H(U^*) - (1/2) \ln(2\pi e(\varepsilon - \varepsilon_{\min})), \quad \varepsilon_{\min} < \varepsilon \leq \varepsilon_{\max}, \quad (2)$$

где $H(U^*)$ — дифференциальная энтропия на множестве оценок U^* с наименьшей средней погрешностью $\varepsilon_{\min} > 0$, $R(\varepsilon_{\min}) \rightarrow \infty$ и $R(\varepsilon_{\max}) = 0$.

Модель классификации данных представлена парой стохастических преобразований $\Omega \rightarrow X \rightarrow \hat{\Omega}$, где Ω — множество меток классов мощности $c \geq 2$, X — множество классифицируемых объектов, $\hat{\Omega}$ — множество оценок классов с мерой погрешности $[\omega \neq \hat{\omega}]$, $\omega \in \Omega$, $\hat{\omega} \in \hat{\Omega}$. Минимум средней взаимной информации $I_{Q_{\hat{\Omega}|X}}(X; \hat{\Omega})$ по распределениям $Q_{\hat{\Omega}|X}$ при ограничении средней погрешности $E_{Q_{\hat{\Omega}|X}}(X; \hat{\Omega}) \leq \varepsilon$ дает монотонно убывающую функцию [1]

$$R(\varepsilon) = I(\Omega; X) - h(\varepsilon - \varepsilon_{\min}^{(c)}) - (\varepsilon - \varepsilon_{\min}^{(c)}) \ln(c - 1), \quad \varepsilon_{\min}^{(c)} \leq \varepsilon \leq \varepsilon_{\max}^{(c)}, \quad (3)$$

где $h(z) = -z \ln z - (1 - z) \ln(1 - z)$, $R(\varepsilon_{\min}^{(c)}) = I(\Omega, X)$ и $R(\varepsilon_{\max}^{(c)}) = 0$.

Обращения функций (1), (2) и (3) дают нижние границы средней погрешности $R^{-1}(I)$ при количестве информации I . Граница в форме (1) остается справедливой для модели сегментации изображений. Граница в форме (2) допускает обобщение для модели оценивания случайного параметра по выборке независимых измерений с заданной плотностью распределения. Существенно, что границы (1)–(3) не зависят от решающих алгоритмов и характеризуют эффективность данных, на которых принимаются решения.

[1] *Lange M. M., Lange A. M.* Information-Theoretic Lower Bounds to Error Probability for the Models of Noisy Discrete Source Coding and Object Classification // *Pattern Recognition and Image Analysis*. — 2022. — Vol. 32, No. 3. — Pp. 570–574. <https://doi.org/10.1134/S105466182203021X>

[2] *Ланге М. М., Ланге А. М.* Нижняя граница погрешности оценивания случайного параметра при заданном количестве информации // *Информатика и ее применения*. — 2024. — Т. 18, № 2. — С. 17–24. <https://doi.org/10.14357/19922264240203>

Кластеризация парных сравнений

*Двоенко Сергей Данилович*¹

sergedv@yandex.ru

¹Тула, Тульский государственный университет

В соответствии с ранее представленной технологией [1], парные сравнения похожести или различий должны быть погружены в многомерное метрическое (евклидово) пространство. Тогда различия рассматриваются как расстояния, а положительные похожести — как скалярные произведения элементов множества (объектов) в одном квадранте пространства.

В этих условиях удается построить полные аналоги известных алгоритмов кластер-анализа и машинного обучения (k -средних, Форель и т. п.) для обработки парных сравнений [2]. Переход от расстояний к скалярным произведениям и наоборот осуществляется на основе известной теоремы косинусов. Именно поэтому можно показать, что результат на основе известного критерия $J(K)$ разбиения объектов на K кластеров по расстояниям до их центров, который минимизируется, в точности соответствует результату на основе критерия разбиения по близостям $I(K) = Const - J(K)$, где критерий $I(K)$ разбиения объектов на K кластеров по близостям с их центрами максимизируется.

Тем не менее, на результат могут повлиять дополнительные преобразования: нормировка расстояний или скалярных произведений. Тогда разбиение по расстояниям может отличаться от разбиения по близостям. Взаимные расстояния между объектами определяют конфигурацию в пространстве и их концентрации на кластеры. Нормировка скалярных произведений уничтожает эту конфигурацию, изменяя взаимные расстояния и располагая объекты в пространстве

на единичной сфере. Но теперь расстояния уже не важны, т. к. концентрации объектов определяются только углами между их векторами.

В работе на примере данных по ирисам [3] показано влияние изменения матриц парных сравнений на результат кластеризации алгоритмом k -средних. В частности, показано, что возможно улучшение разбиения пересекающихся второго и третьего кластеров по сравнению с известными результатами.

Работа поддержана грантом РФФИ № 20-07-00055.

- [1] *Двоенко С. Д.* Технология кластеризации парных сравнений // ММРО-21. Тез. докл., Москва, 2023 г. — М.: РАН, 2023. — С. 31.
- [2] *Dvoenko S.* Clustering of Data Represented by Pairwise Comparisons // Control and Cybernetics. — 2022. — Vol. 51, No. 3. — P. 343–387. <https://doi.org/10.2478/candc-2022-0021>
- [3] *Fisher R A.* The Use of Multiple Measurements in Taxonomic Problems // Ann. Eugenics. — 1936. — Vol. 7, No. 2. — P. 179–188.

Дистилляция знаний в применении к моделям бустинга

*Неделько Виктор Михайлович*¹

`nedelko@math.nsc.ru`

¹Новосибирск, Институт математики им. С. Л. Соболева

Дистилляция знаний — активно используемый подход в машинном обучении, который заключается в том, что при обучении некоторой модели вместо исходных данных (или в дополнение к ним) используется выход от уже обученной ранее модели. Будем рассматривать случай, когда сложность модели-учителя значительно выше, чем сложность модели-ученика. Этот случай наиболее типичен, поскольку дистилляция изначально придумана именно как метод упрощения модели.

Известно, что обучение модели на прогнозах модели-учителя зачастую позволяет достичь большей точности, чем при непосредственном обучении малой модели на исходных данных. Этот факт контринтуитивен. Действительно, если модель-ученик способна давать решения с такой же точностью, как и модель-учитель, то ее сложность, как минимум, достаточная (для описания данных). Значит сложность (гибкость) модели-учителя избыточна. Однако, как известно, избыточная гибкость модели может привести к переобучению, поэтому приходится подбирать момент останова. В целом, остается непонятным, за счет чего модели избыточной сложности могут работать лучше моделей меньшей, но заведомо достаточной сложности.

Методы дистилляции, как правило, применяются к нейросетевым моделям. Однако этот подход оказывается возможным применить и для градиентного бустинга.

Модель бустинга имеет два основных параметра, определяющих ее сложность: размер решающего дерева и количество деревьев. При этом за гиб-

кость модели (способность отражать сложные зависимости переменных) отвечает именно число вершин дерева [1].

В ходе данной работы был проведен следующий численный эксперимент. Была построена некоторая вероятностная модель для генерации данных в задаче классификации.

Для выбранной модели оптимальная (байесовская) решающая функция может быть сколь угодно точно приближена ансамблем деревьев глубины 2. Вместе с тем, при обучении по выборке лучшая точность была достигнута для бустинга на деревьях глубины 7.

Далее была реализована идея дистилляции, а именно: сгенерирована размеченная выборка большего размера и на ней сделан прогноз обученной ранее моделью глубины 7. После этого на спрогнозированных значениях была обучена модель глубины 2. Для всех моделей проводился подбор оптимального числа деревьев.

Оказалось, что модель бустинга, полученная методом дистилляции, имеет точность, значительно превышающую точность модели глубины 2, обученной на исходной выборке. Подобный эффект наблюдался при разных размерах обучающей выборки и разных моделях.

Таким образом, идея дистилляции может применяться в том числе к методу бустинга, что дает дополнительный материал для понимания механизмов ее эффективности.

Работа выполнена в рамках госзадания Института математики им. С. Л. Соболева (проект № FWNF-2022-0015).

- [1] *Неделько В. М.* К вопросу об эффективности бустинга в задаче классификации // Вестник Новосибирского государственного университета. Серия: Математика, механика, информатика. — 2015. — Т. 15, № 2. — С. 72–89. <https://doi.org/10.17377/PAM.2015.15.206>

Система распознавания на основе общности свойств

*Краснопрошин Виктор Владимирович*¹

krasoproshin@bsu.by

*Карканица Анна Викторовна*²

a.karkanica@grsu.by

Родченко Вадим Григорьевич^{2*}

rovar@grsu.by

¹Минск, Белорусский государственный университет

²Гродно, Гродненский государственный университет имени Янки Купалы

В информатике проблема обучения распознаванию образов относится к числу фундаментальных и изучается в теории интеллектуальных систем [1]. Обучение основано на анализе частных эмпирических данных, задаваемых прецедентами, и выполняется в рамках решения задачи распознавания [2].

В машинном обучении сложился и абсолютно доминирует традиционный подход (*определенный стандарт*), который предусматривает, что процесс обу-

чения базируется на основе использования принципа «разделяй и классифицируй», когда фактически образ каждого класса задается множеством разрозненных объектов в признаковом пространстве, и затем решается оптимизационная задача построения разделяющей поверхности [3].

Предлагается альтернативный подход к обучению, который базируется на исследовании свойств классов и выявлении таких, которые обеспечивают их различие. Концептуально процесс обучения реализуется на основе результатов построения областей определения классов (*образов классов*) и сравнения их взаимного размещения в рамках «сравнения целого с целым».

Предположим, что имеются множество описаний объектов X , множество допустимых ответов Y , обучающая выборка $X^m = \{(x_1, y_1), \dots, (x_m, y_m)\}$, сформированная на основе словаря признаков $F = \{f_1, \dots, f_n\}$. Обозначим через $V = \{v_1, \dots, v_q\}$ множество всех непустых подмножеств, образованных всевозможными сочетаниями признаков из F , где $q = \sum_{i=1}^n C_n^i = 2^n - 1$.

Выявление среди множества $V = \{v_1, \dots, v_q\}$ признаковых подпространств, в которых образы классов не пересекаются проводится следующим образом:

1. выбираем очередное сочетание v_i , $i = \overline{1, q}$, и на основе всех его признаков определяем соответствующее признаковое подпространство;
2. в этом признаковом подпространстве строим образы классов (в виде кластерных структур) и проводим оценку их взаимного размещения [4];
3. сочетание признаков v_i включаем в результирующее множество V^* только тогда, когда образы классов не пересекаются.

В результате анализа всех элементов $V = \{v_1, \dots, v_q\}$ будет построено множество $V^* = \{v_1^*, \dots, v_t^*\}$, где $0 \leq t \leq q$, в котором каждому отдельному подмножеству признаков $v_i^* \in V^*$ соответствует выявленная закономерность: «в пространстве признаков подмножества v_i^* классы не пересекаются». Отметим, что сочетание признаков $v_i^* \in V^*$ определяет пространство решений, в котором классы не пересекаются, т.е. для образов классов внутри таких пространств подтверждается гипотеза компактности, и построение алгоритма классификации превращается в тривиальную процедуру.

Предложенный подход успешно применялся для обработки реальных наборов данных в рамках решения прикладных задач и для проведения численных экспериментов на основе модельных данных [5].

- [1] *Абламейко С. В., Краснопрошин В. В., Образцов В. А.* Модели и технологии распознавания образов с приложением в интеллектуальном анализе данных // Вестник БГУ. Сер. 1, Физика. Математика. Информатика. — Минск: БГУ, 2011. — № 3. — С. 62–72.
- [2] *Ту Дж., Гонсалес Р.* Принципы распознавания образов // М.: Мир, 1978. — 412 с.
- [3] *Краснопрошин В. В., Родченко В. Г.* Обучение по прецедентам на основе анализа свойств признаков // Доклады БГУИР. — Минск: БГУИР, 2017. — № 6. — С. 35–41.

- [4] *Краснопрошин В. В., Родченко В. Г.* Кластерные структуры и их применение в интеллектуальном анализе данных // Информатика. — Минск: ОИПИ НАН Беларуси, 2016. — № 2. — С. 71–77.
- [5] *Rodchenko V.* Automatic Detection of Hidden Regularities Based on the Study of Class Properties // Pattern Recognition and Image Analysis. — 2020. — No. 30, — С. 222–227.

Построение и исследование логических корректоров над произведением частичных порядков

Генрихов Игорь Евгеньевич^{1*}

ingvar1485@rambler.ru

*Дюкова Елена Всеволодовна*¹

edjukova@mail.ru

¹Москва, ФИЦ ИУ РАН

Рассматривается логический подход к задаче классификации по прецедентам с непересекающимися классами и целочисленными признаками. При конструировании логических классификаторов большое внимание уделяется вопросам синтеза корректных алгоритмов, т. е. алгоритмов, не ошибающихся на обучающей выборке. Базовым понятием является понятие элементарного классификатора (ЭК). ЭК — это элементарная конъюнкция, определенная на признаковых описаниях прецедентов. Если на описании прецедента S элементарная конъюнкция обращается в единицу, то говорят, что S содержит данный ЭК. ЭК называется корректным для класса K , если не существует двух прецедентов таких, что один из них принадлежит K , а второй не принадлежит K , и оба содержат данный ЭК.

В классических моделях в процессе обучения строятся семейства корректных ЭК. На этапе классификации каждый найденный ЭК участвует в процедуре голосования и формирования оценок принадлежности распознаваемого объекта к классам. Трудными являются задачи, в которых практически каждый корректный ЭК содержится в небольшом числе прецедентов. Подобная ситуация возникает, например, в случае, когда признаки имеют много значений.

В [1] предложены модели логических корректоров, в которых корректность классификации достигается путем использования не обязательно корректных ЭК и построения семейств корректных наборов ЭК. Набор U , состоящий из ЭК, называется корректным для класса K , если для любой пары обучающих объектов, один из которых принадлежит K а другой не принадлежит K , в наборе U существует ЭК, принимающий разные значения на этих объектах. Если такой ЭК обращается в 1 на объекте из K , то набор U называется монотонным корректным. Основная проблема — вычислительная сложность этапа обучения, на котором решается задача поиска минимальных покрытий булевой матрицы L значительных размеров. В [2] построена модель с ЭК ранга 1 и с применением генетического алгоритма для поиска минимальных покрытий матрицы L .

В настоящей работе предложено поиск монотонных корректных наборов из ЭК ранга 1 осуществлять на основе построения минимальных покрытий булевой матрицы с числом столбцов равным числу признаков, что позволяет существенно сократить временные затраты. Рассмотрены вопросы разработки стохастических моделей логических корректоров и модификации этих моделей на случай, когда на множествах значений признаков заданы конечные частичные порядки. Построение классических логических классификаторов над произведением частичных порядков проведено в [3].

Исследование выполнено за счет гранта Российского научного фонда № 24-21-00301, <https://rscf.ru/project/24-21-00301/>.

- [1] *Дюкова Е. В., Журавлёв Ю. И., Рудаков К. В.* Об алгебраическом синтезе корректирующих процедур распознавания на базе элементарных алгоритмов // Ж. вычисл. мат. и матем. физ. — 1996. — Т. 36, № 8. — С. 217–225.
- [2] *Djukova E. V., Zhuravlev Yu. I., Sotnezov R. M.* Construction of an Ensemble of Logical Correctors on the Basis of Elementary Classifiers // Pattern Recognition and Image Analysis. — 2011. — Vol. 21. No. 4. — Pp. 599–605.
- [3] *Дюкова Е. В., Масляков Г. О., Прокофьев П. А.* О логическом анализе данных с частичными порядками в задаче классификации по прецедентам // Ж. вычисл. мат. и матем. физ. — 2019. — Т. 59, № 9. — С. 1605–1616.

Методы топологического анализа для порождения более информативных синтетических признаков на основе опорных цепей и произвольных метрических функций расстояния

*Торшин Иван Юрьевич*¹

tiy135@yahoo.com

¹Москва, ФИЦ ИУ РАН

Топологическая теория распознавания является развитием алгебраического подхода к распознаванию научной школы Ю. И. Журавлёва в применении к плохо формализованным задачам распознавания, классификации, прогнозирования. Одной из целей данной теории, осуществляющей переход от множества описаний объектов X к топологии $T(X)$ и решетке $L(T(X))$, является разработка методов систематического порождения и отбора синтетических признаков описаний объектов, которые бы характеризовались большей «информативностью» по отношению к поставленным задачам, чем исходные признаки.

В настоящей работе прецедентные соотношения моделируются как соответствующие массивы расстояний в том или ином метрическом пространстве, порождаемом над решеткой $L(T(X))$. Исследованы способы вычисления таких расстояний и методы порождения синтетических признаков описаний. Анализ формализма топологической теории распознавания на основе фундаментальных понятий функционального анализа позволил предложить ранее неисследованные подходы к определению решеточных оценок — в частности, исполь-

зование опорных цепей для анализа $L(T(X))$. Для практического приложения формализма предложены подходы к исследованию свойств, способы оценки релевантности функций по отношению к решаемым задачам, способы порождения и отбора синтетических признаков.

Представлены результаты экспериментальной апробации алгоритмов топологического анализа данных на задачах фармакоинформатики (анализ хемокиномных данных). Материал обучения/тестирования включил 2400 выборок данных «молекула — свойство» из ProteomicsDB, в качестве свойств молекул рассматривались константы ингибирования (EC_{50}), активности киназ для различных концентраций ($E_j(C_i)$). Исходные признаки определялись как булевы инварианты над множествами хи-цепей и хи-узлов хемографов.

Наилучший результат применения нового «топологического» формализма с полиномиальным корректором (коэффициент ранговой корреляции на контроле $rc = 0,90 \pm 0,23$) немного превзошел наилучший результат применения метода опорных функций ($rc = 0,86 \pm 0,20$). Нейросетевые корректоры, вне зависимости от использованных конфигураций, отличались крайне низкими показателями ($r = 0,45 \pm 0,22$, $rc = 0,22 \pm 0,21$), а «случайный лес» приводил к существенному переобучению: в 12% выборок данных ($n = 290$) корректор «случайный лес» приводил к улучшению результатов по сравнению с наилучшими полиномиальными корректорами, а в 70% выборок данных ($n = 1670$) — к ухудшению.

Работа выполнена при поддержке гранта РФФИ (проект 23-21-00154).

- [1] *Торшин И. Ю.* О формировании множеств прецедентов на основе таблиц разнородных признаковых описаний методами топологической теории анализа данных // Информатика и ее применения. — 2023. — Т. 17, № 7. — С. 2–10.

О логической классификации частично упорядоченных целочисленных данных

*Дюкова Елена Всеволодовна*¹

edjukova@mail.ru

Дюкова Анастасия Петровна^{1*}

anastasia.d.95@gmail.com

¹Москва, ФИЦ ИУ РАН

Рассмотрены актуальные вопросы анализа частично упорядоченных целочисленных данных, возникающие в машинном обучении при синтезе логических классификаторов.

В [1, 2] на основе обобщения базовых понятий усовершенствован традиционный подход к построению логических процедур классификации по прецедентам. Впервые построены классификаторы, ориентированные на задание частичных порядков на множествах значений признаков (в этом случае описания прецедентов — элементы декартова произведения конечных частично упорядоченных множеств). Создание таких классификаторов обусловлено существованием

прикладных задач, качественное решение которых невозможно в рамках классического подхода, когда множество значений каждого признака — антицепь. Исследована зависимость качества логической классификации от выбора частичных порядков на множествах значений признаков. В настоящее время указанные построения и исследования проведены для трех основных направлений логической классификации, а именно: Correct Voting Procedures или CVP (предложено в отечественных научных школах С. В. Яблонского и Ю. И. Журавлёва), Logical Analysis of Data [3] и направления, основанного на идеях ДСМ-метода В. К. Финна [4]. Усовершенствованные модели классификаторов двух последних направлений описаны соответственно в работах [5] и [6].

В [7] создана общая схема работы моделей логических алгоритмов классификации по прецедентам над произведением антицепей. В единой терминологии приведено оригинальное описание названных выше основных направлений логической классификации. Согласно предложенной схеме каждый классификатор на этапе обучения задает некоторый частичный порядок на специальном множестве фрагментов описаний прецедентов, называемых в CVP представительными элементарными классификаторами, и ищет максимальные относительно заданного порядка элементы этого множества.

Кроме обзора перечисленных выше результатов в докладе представлены новые исследования. Поставлена и решена важная в методологическом плане задача создания единой схемы обучения логического классификатора над декартовым произведением конечных частичных порядков. При описании этой схемы использована терминология обобщенных на случай частично упорядоченных данных процедур CVP. В качестве базового понятия выступает понятие представительного элементарного классификатора над декартовым произведением конечных частичных порядков, введенное в [1].

Исследование выполнено за счет гранта Российского научного фонда № 24-21-00301, <https://rscf.ru/project/24-21-00301/>.

- [1] Дюкова Е. В., Масляков Г. О., Прокофьев П. А. О логическом анализе данных с частичными порядками в задаче классификации по прецедентам // Ж. вычисл. матем. и матем. физ. — 2019. — Т. 59, № 9. — С. 1605–1616.
- [2] Дюкова Е. В., Масляков Г. О. О выборе частичных порядков на множествах значений признаков в задаче классификации // Информатика и ее применения. — 2021. — Т. 15, № 4. — С. 74–80.
- [3] Hammer P. L. Partially Defined Boolean Functions and Cause-Effect Relationships // In: Lecture at the Int. Conf. Multi-Attribute Decision Making Via ORBased Expert Systems. — University of Passau, Passau, Germany, 1986.
- [4] Финн В. К. Базы данных с неполной информацией и новый метод автоматического порождения гипотез // В кн.: Диалоговые и фактографические системы информационного обеспечения. — М.: ВСНТО, 1981. — С. 153–156.
- [5] Anisimova D., Djukova E., Djukova A. Supervised Classification Problem: Searching for Maximum Patterns // X International Conference on Information Technology

and Nanotechnology (ITNT), Samara, Russian Federation, 20–24 мая 2024. — NJ: IEEE, 2024.

- [6] Дюкова Е. В., Масляков Г. О., Янаков Д. С. Корректная классификация по прецедентам: ДСМ-метод над произведением частичных порядков // Информатика и её применения. — 2024. — Т. 18, № 3. — С. 56–63. В печати.
- [7] Дюкова Е. В., Масляков Г. О., Дюкова А. П. Логические методы корректной классификации данных // Информатика и её применения. — 2023. — Т. 17, № 3. — С. 64–70.

Новые методы построения оптимальных ансамблей

*Сенько Олег Валентинович*¹*

senkoov@mail.ru

*Докужин Александр Александрович*¹

dalex@ccas.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

Ансамблевые методы продолжают играть существенную роль в машинном обучении, демонстрируя во многих случаях обобщающую способность, превосходящую обобщающую способность альтернативных технологий. В настоящее время два главных направления ансамблевых методов — случайные леса и градиентный бустинг.

В методе случайный лес ансамбль деревьев строится по обучающим выборкам, которые генерируются из исходной выборки с помощью процедур бэггинга и метода случайных подпространств. Каждое дерево строится исходя из условия максимальной точности аппроксимации, а коллективное решение вычисляется как среднее по прогнозам, вычисляемым отдельными деревьями ансамбля.

В методе градиентный бустинг коллективное решение ищется как линейная комбинация деревьев, наилучшим образом аппроксимирующая зависимость целевой переменной от признаков. На каждом шаге в линейную комбинацию добавляется новое слагаемое исходя из цели минимизации потерь.

Новый способ построения ансамблей был ранее предложен в работах авторов, в т. ч. [1]. Данный способ нацелен на построение такого ансамбля, чтобы коллективное решение, вычисляющее средний прогноз по ансамблю, обладало бы наилучшей обобщающей способностью.

Было показано, что ошибка усредненного прогноза может быть представлена как разность двух составляющих: среднего от математических ожиданий квадратичных ошибок алгоритмов, вошедших в ансамбль; среднего от математических ожиданий квадратов разностей между прогнозами, вычисляемых алгоритмами, вошедшими в ансамбль. Таким образом, для достижения высокой обобщающей способности необходимо одновременно как высокая обобщающая способность отдельных алгоритмов ансамбля, так и значительное расхождение вычисляемых ими прогнозов.

В настоящей работе предложен вычислительно эффективный метод, позволяющий достигнуть выполнения двух упомянутых условий. В данном методе на каждом шаге k в текущий ансамбль E_{k-1} добавляется алгоритм A , для которого достигается минимум специального функционала $Err(A) - \mu(A - E_k)^2$, где $E_k = ((k - 1)E_{k-1} + A)/k$.

Эксперименты, которые проводились с использованием в качестве элементов ансамбля отдельных деревьев а также линейных комбинаций деревьев, получаемых с помощью градиентного бустинга, подтвердили эффективность подхода.

- [1] *Докукин О. В., Сенько О. В.* Новый двухуровневый метод машинного обучения для оценивания вещественных характеристик объектов // Известия РАН. Теория и системы управления. — 2023. — № 4. — С. 17–24.

Гармонические меры качества для многоклассовой классификации

Копылов Андрей Валериевич^{1*}

and.kopylov@gmail.com

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

¹Тула, Тульский государственный университет

Анализ литературы [1] показывает растущую популярность использования F -меры для оценки качества в современных исследованиях машинного обучения. В данной работе рассматривается задача оценивания качества многоклассовой классификации, которая сводится, как правило, к применению метрик двухклассовой (бинарной) классификации по схемам «один от всех» (One-Vs-The-Rest, OVR) или «каждый от каждого» (One-Vs-One, OVO) с тем или иным способом усреднения оценок [2]. Поскольку OVO классификация на K классов порождает по крайней мере $K(K + 1)/2$ двухклассовых задач, в то время как OVR лишь K мы рассматриваем последнюю из стратегий, хотя те же соображения могут быть применены и в случае OVO.

Классификация на K классов по схеме один от всех порождает K двухклассовых матриц спутанности. В результате получим K оценок F -меры в зависимости от того, какой класс выбран в качестве целевого (позитивного).

Реализованная в широко распространенных библиотеках машинного обучения мера F_{macro} вычисляется как арифметическое среднее бинарных оценок. Хотя усреднение мер F по схеме многоклассового OVR частично устраняет их отрицательные свойства, мера F_{macro} часто дает более оптимистичную оценку, чем того требует ситуация [1, 3].

В отличие от арифметического усреднения F_{macro} мы предлагаем применить гармоническое усреднение оценок соответствующих классов, сохранив в свою

очередь исходную идею F -меры как гармонического среднего:

$$F_{harm} = \frac{K}{\sum_{k=1}^K \frac{1}{F_k}}.$$

Чтобы избежать известных недостатков F -меры, связанных с тем, что она не учитывает количество верно распознанных объектов негативного класса [1], предлагается также использовать меру P_4 , описанную в [3, 4], которая основана на гармоническом усреднении четырех значений, два из которых являются значениями точности и полноты для случая, когда первый класс назван позитивным, а другие два — значениями точности и полноты для случая, когда позитивным назван другой класс. Данная мера обладает хорошей интерпретируемостью и не требует сложных вычислений. В итоге получим меру качества следующего вида для многоклассовых задач:

$$P_{4harm} = \frac{K}{\sum_{k=1}^K \frac{1}{P_{4(k)}}}.$$

Обе предложенные гармонические меры демонстрируют повышенную чувствительность к крайним случаям несбалансированности данных или нерегулярности правил принятия решений по сравнению с мерой F_{macro} . Например, значения обеих мер будут равны нулю, если объекты хотя бы одного класса не распознаются вовсе. Таким образом, эти меры удобно применять, когда очень важно исключить слабые классификации. Кроме того, учет верно распознанных объектов негативного класса в P_{4harm} делает эту меру чувствительной к дисбалансу классов, в отличие от F_{harm} . Экспериментальные исследования, как на синтетических, так и реальных данных подтверждают описанные выше уникальные свойства предложенных мер с точки зрения задач классификации.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания FEWG-2024-0001.

- [1] *Christen P., Hand D. J., Kirielle N.* A Review of the F-Measure: Its History, Properties, Criticism, and Alternatives // ACM Computing Surveys. — 2023. — Vol. 56, No. 3, Article No. 73. — Pp. 1–24.
- [2] *Farhadpour S., Warner T. A., Maxwell A. E.* Selecting and Interpreting Multiclass Loss and Accuracy Assessment Metrics for Classifications with Class Imbalance: Guidance and Best Practices // Remote Sensing. — 2024. — Vol. 16, No. 3. — P. 533.
- [3] *Копылов А. В., Середин О. С.* Новая F-мера качества бинарных классификаторов // Математические методы распознавания образов: Тезисы докладов 21-й Всероссийской конференции с международным участием, г. Москва 2023 г. — М.: Российская академия наук, 2023. — С. 48–50.
- [4] *Sitarz M* Extending F1 Metric, Probabilistic Approach // arXiv. — 2022. — <https://arxiv.org/abs/2210.11997>

Повышение точности прогнозов в машинном обучении с использованием двухуровневого ансамбля решающих деревьев

Картавецев Антон Александрович¹★

kartavcevaa@my.msu.ru

¹Москва, Факультет ВМК МГУ имени М. В. Ломоносова

Ансамблевые методы машинного обучения, такие как случайные леса и градиентный бустинг, широко используются для решения задач регрессии и классификации. В банковской сфере даже незначительное улучшение точности моделей может привести к существенному увеличению доходов, что определяет важность разработки более эффективных методов. Стандартные методы построения ансамблей не всегда обеспечивают оптимальные результаты, особенно в условиях высокой корреляции между моделями и риска переобучения.

В данной работе проводится исследование трех новых методов — дивергентного леса [1], пересекающихся деревьев (авторский метод) и нежадных деревьев [2] — с целью улучшения качества прогнозов и сравнения их с существующими методами на примере данных из банковской сферы. Дивергентный лес направлен на создание разнообразных деревьев решений с минимальной корреляцией между ними. Пересекающиеся деревья оптимизируют структуру деревьев решений за счет частичного пересечения подмножеств данных, что снижает вероятность переобучения и ускоряет расчет. Нежадные деревья строят деревья решений с учетом множества путей разбиения данных на каждом шаге, что позволяет находить более оптимальные решения.

В настоящее время проводится экспериментальное сравнение предложенных методов с популярными ансамблевыми алгоритмами, такими как градиентный бустинг и случайный лес, на датасетах, включая данные из банковской сферы. Предполагается, что новые методы смогут превзойти существующие в задачах регрессии и классификации, демонстрируя улучшенную точность. Предложенные методы — дивергентный лес, пересекающиеся деревья и нежадные деревья — имеют потенциал улучшить качество прогнозов в различных задачах, особенно в области банковского анализа данных. Их применение может привести к значительному повышению эффективности моделей машинного обучения и, соответственно, к увеличению доходности бизнеса.

- [1] *Sen'ko O. V., Dokukin A. A., Kiselyova N. N., Dudarev V. A., Kuznetsova Yu. O.* New Two-Level Ensemble Method and Its Application to Chemical Compounds Properties Prediction // *Lobachevskii J. Mathematics.* — 2023. — No. 44. — Pp. 188–197.
- [2] *Norouzi M., Collins M., Johnson M. A., Fleet D. J., Kohli P.* Efficient Non-Greedy Optimization of Decision Trees // *Advances in Neural Information Processing Systems.* — 2015.

Generating Human Object Interaction via Neural Ordinary Differential Equations for Zero-shot Robot Manipulation

Yang Chen^{1*}

chenyang@smbu.edu.cn

Wenrui Wang²

wangwenrui@stu.hit.edu.cn

¹Shenzhen, Faculty of Computational Mathematics and Cybernetics, Shenzhen MSU-BIT University

²Harbin, College of Computer Science and Technology, Harbin Institute of Technology

This paper aims to achieve zero-shot manipulation by translating human-object interaction to robot manipulation, to alleviate dependence on manipulation data. We separate this task into two sub-tasks: plan prediction and action generation. In plan prediction, we propose the Neural ODE plan prediction model (ODE-Plan) based on neural ordinary differential equations (NODEs) to generate masked human-object interaction series given an initial image and a goal image. In action generation, we propose the robot action generation model (RAG) to translate the masked interaction plan into actual robot action.

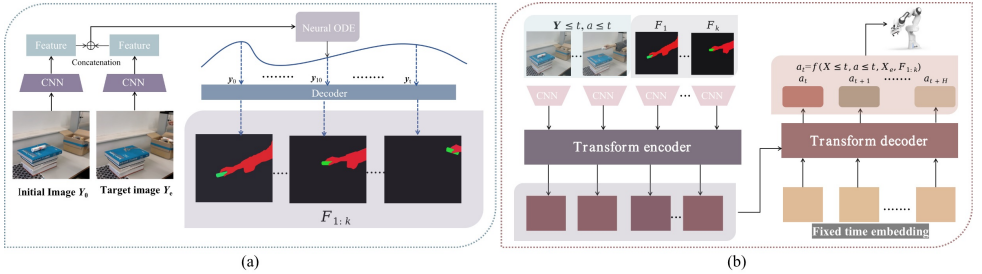


Fig. 1. (a) The architecture of ODE-Plan (b) The architecture of RAG

Fig.1(a) shows the architecture of ODE-Plan. Neural ODEs [1] provide a novel framework for designing and training flexible, continuous-time neural network architectures. In the real scene, given an initial image of the action $\mathbf{Y}_0 \in \mathbb{R}^{H \times W \times 3}$ and an end image of the action $\mathbf{Y}_e \in \mathbb{R}^{H \times W \times 3}$, where H and W are the height and width of the image, Neural ODE model generates plausible future masks $\mathbf{F}_{1:K} \in \mathbb{R}^{H \times W}$ that capture the anticipated trajectories of hands and objects, as:

$$\mathbf{z}(0) = \text{Conv}(\mathbf{Y}_0) \oplus \text{Conv}(\mathbf{Y}_e), \quad \frac{d\mathbf{z}}{dt}(t) = f_\theta(t, \mathbf{z}(t)), \quad (1)$$

where $\text{Conv}(\cdot)$ is the convolutional network, \oplus is concatenation, $f_\theta(\cdot)$ is a linear layer, and $\mathbf{z}(t) \in \mathbb{R}^D$ is the latent feature to model the interaction dynamics at time step t . From (1), the state $\mathbf{z}(t)$ of such a system is defined at any time T , and can be computed at any desired time using a numerical ODE solver, which will evaluate

the dynamics f_θ to determine the solution, as:

$$\mathbf{z}(T) = \text{ODESolver}(f_\theta, \mathbf{z}(0), T) \approx \mathbf{z}(0) + \int_0^T f_\theta(t, \mathbf{z}(t)) dt. \quad (2)$$

Then we can use an image decoder to map latent feature to the masked interaction plan $\mathbf{F}_{1:K}$, as $\mathbf{F}_k = \text{decoder}(\mathbf{z}(k))$.

Fig. 1(b) shows the architecture of RAG. We leverage a transformer [2] that is conditioned on the outputs of the ODE-Plan $\mathbf{F}_{1:K}$ and for each observation \mathbf{Y}_t , and predicts actions $\mathbf{a}_{t:t+H}$ for H steps in the future. This instantiates as a closed-loop policy $\pi(\mathbf{a}_t | \mathbf{F}_{1:K}, \mathbf{Y}_t, \mathbf{Y}_e)$, which executes the first predicted action and obtains the current observation for the next loop.

This research has received funding from Shenzhen BSQD project under Grant Agreement No. RCBS20231211090749086.

- [1] *Chen R. T. Q., Rubanova Yu., Bettencourt J., Duvenaud D. K.* Neural ordinary differential equations // Advances in Neural Information Processing Systems. — 2018. — Vol. 31.
- [2] *Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I.* Attention is all you need // Advances in Neural Information Processing Systems. — 2017. — Vol. 30.

Methods for Enhancing Intelligent Type Inference for the Python Language

Valeriya Abdrakhmanova^{1*}

vabdrahmanova@ispras.ru

*Konstantin Sorokin*¹

ksorokin@ispras.ru

¹Moscow, ISP RAS

Currently, there is a growing interest in the automatic prediction of missing type annotations in programs written in dynamically typed programming languages. Type inference methods play a crucial role in source code analysis by automatically identifying variable and function types, which enhances code comprehension and facilitates maintenance and development. One of the most effective intelligent approaches utilizing small-sized models for type inference in Python to date is the TypeT5 [1] approach. This method leverages the CodeT5 pretrained language model, treating type prediction as a code infilling task.

The main goal of this work is to showcase limitations of the TypeT5 approach and develop methods for improving it. This includes the usage of more recent base model architectures, enhancement of the model input context, and optimization of the inference.

An analysis of existing type inference solutions was conducted, with a detailed study of the TypeT5 approach and its components to identify shortcomings. The

identified shortcomings of the approach include: the presence of redundant type predictions, which negatively affect performance, the incompleteness of the model's input data, the lack of handling for stub files, and the use of an outdated model. To improve the accuracy of type predictions, modern models that outperform CodeT5 [2] in code completion tasks were employed. The model's input data was supplemented with initially missing information about external dependencies. Type inference optimizations were suggested by eliminating redundant predictions and modifying static code analysis pipeline for context construction, ensuring that model inputs incorporate information from stub files.

The developed methods resulted in a 5–9% increase in accuracy and up to a 10% reduction in time. Furthermore, the TypeT5-based type inference tool was adapted for integration with the Gerrit code review system.

The methods and approaches proposed in this work are of interest to software developers and researchers in the field of information technology.

- [1] *Wei J., Durrett G., Dillig I.* TypeT5: Seq2seq Type Inference using Static Analysis // International Conference on Learning Representations. — 2023.
- [2] *Wang Y., Wang W., Joty Sh., Hoi S. C. H.*, CodeT5: Identifier-aware Unified Pre-trained Encoder-Decoder Models for Code Understanding and Generation // Proc. 2021 Conf. Empirical Methods in Natural Language Processing. — 2021. — Pp. 8696–8708.

Повышение структурной согласованности в задаче поиска групп идентичных объектов

Рябцев Антон Борисович^{1*}

ryabtsev.ab@phystech.edu

*Дуллин Сергей Константинович*¹

skdulin@mail.ru

¹Москва, МФТИ (НИУ)

В современном мире электронной торговли ежедневно появляются миллионы товаров на различных маркетплейсах, предоставляя покупателям огромный выбор и разнообразие. Однако, в контексте такого разнообразия предложений, возникает вопрос интероперабельности — способности двух или более информационных компонентов к обмену информацией и использованию информации, полученной в результате этого обмена [1]. На практике далеко не все товарные предложения на маркетплейсе уникальны. Существуют карточки товаров, заведенные разными продавцами и отличающиеся по содержанию, но идентичные семантически. При этом одни продавцы предлагают более выгодную цену, другие менее выгодную. Склейка карточек идентичных товаров играет важную роль в улучшении пользовательского опыта покупателей, в то же время склейка карточек неидентичных (т. е. не обладающих интероперабельностью) товаров, напротив, приводит к ухудшению пользовательского опыта.

Самый очевидный подход [2] к склейке товаров основан на принципе транзитивности. Принцип реализует логику доклеивания нового товара к готовой группе, если он схож хотя бы с одним товаром этой группы. Однако такой метод, хоть и прост в реализации, порождает серьезные проблемы, нарушая интероперабельность создаваемых групп. Вместо четко выраженных кластеров карточек идентичных товаров (каждый элемент имеет положительную связь с каждым), типично явление рассогласованности внутри групп, при котором связь между некоторыми элементами группы неоднозначна.

В [3] было рассмотрено влияние структуры взаимодействующих элементов на мотивацию к интероперабельности на основе определенных характеристик в структуре элементов. Идеальная ситуация (консонанс) — когда между всеми элементами одной группы установлены положительные связи, но со всеми другими элементами связи отрицательные. На практике же типична ассонансная структура системы, поэтому ставится задача поиска ближайшего к ассонансной структуре консонансного прообраза, алгоритм решения которой описан в [3]. Однако, поиск такого прообраза даже для поликонсонанса степени 2 является сложной комбинаторной задачей. Другим подходом к поиску близкого консонансного прообраза, не требующим решения тяжелых комбинаторных задач, является алгоритм распространения меток [4]. В основе алгоритма лежит идея о том, что у вершины графа сообщество будет таким же, как и у большинства ее соседей.

Для сравнения алгоритма из [4] с подходом из [2] использовался набор реальных товаров с маркетплейса. Была выбрана 1 категория товаров, состоящая из порядка 20 000 товаров. Из них случайным образом было выбрано 3000 товаров, которые были разбиты на группы вручную. В результате получилось 704 группы. Для оценки качества полученного решения использовался Rand Index [5]:

$$RI = \frac{TP + TN}{TP + FP + TN + FN}.$$

Для сравнения подходов использовался усредненный по всем 20 категориям RI, результаты приведены в Табл. 1.

Таблица 1. Rand Index двух сравниваемых алгоритмов склейки одинаковых товаров

Транзитивное замыкание	Алгоритм распространения меток
0,696	0,863

- [1] *ГОСТ Р. 55062-2012* Информационные технологии (ИТ). Системы промышленной автоматизации и их интеграция. Интероперабельность. Основные положения. — М.: Стандартинформ. — 2014.
- [2] *Баас Ж., Дастан М. М., Филдерс А. Ж.* Использование транзитивности для сопоставления сущностей // ESWC Сателлитные мероприятия: виртуальное мероприятие, пересмотренные избранные статьи. — 2021. — С. 109–114.

- [3] *Розенберг И. Н., Дулин С. К., Дулина Н. Г.* Моделирование структуры интероперабельности средствами структурной согласованности // Информатика и ее применения. — 2023. — Т. 17, № 1. — С. 57–65.
- [4] *Жу С., Гахрамани З.* Обучение на основе размеченных и неразмеченных данных с помощью распространения меток // Технический отчет, Университет Карнеги-Меллона. — 2002. [Электронный ресурс] — <https://github.com/snknitin/Learning-from-Labeled-and-Unlabeled-Data-with-Label-Propagation/blob/master/label-propagation.pdf> (дата обращения: 20.02.2024)
- [5] *Рэнд В. М.* Объективные критерии оценки методов кластеризации // Журнал Американской статистической ассоциации. — 1971. — Т. 66, № 336. — С. 846–850.

Применение бустинга в рекомендательных системах

Жарова Мария Александровна^{1,2,*}

zharova.ma@phystech.edu

*Цурков Владимир Иванович*²

v.tsurkov@frccsc.ru

¹Москва, Московский физико-технический институт (национальный исследовательский университет)

²Москва, Федеральный исследовательский центр «Информатика и управление» РАН

В современных рекомендательных системах широко используется гибридный подход, объединяющий модели матричной факторизации и контентные алгоритмы. Комбинация этих методов актуальна, т. к. коллаборативная фильтрация обеспечивает точные и быстрые прогнозы [1], а контентные методы устраняют проблему холодного старта и уточняют рекомендации на основе метаданных [2].

Однако, при ограниченном числе объектов методы коллаборативной фильтрации становятся менее эффективными из-за уменьшения разреженности матриц и появления большого количества дубликатов векторов пользователей [3]. К этому добавляется проблема высокой корреляции некоторых объектов друг с другом, поэтому контентный алгоритм в отдельности не подойдет для решения поставленной задачи.

В данной работе предлагается построить систему из независимых классификационных моделей для каждого объекта, обучение каждой из которых базируется на метаданных и агрегированных коллаборативных факторах. Далее при помощи калибровки предсказаний можно получить сравнимые вероятности заинтересованности пользователей в различных объектах, из которых будет формироваться итоговый список рекомендаций [4].

В качестве обучающих данных для каждой модели используется информация о 500 тыс. пользователях и 10 уникальных объектах. Результаты моделирования по метрике ROC-AUC для тренировочной, валидационной и тестовой выборок представлены на Рис. 1. Для обучения выбрана модель CatBoost [5] с подобранными гиперпараметрами, для калибровки применяется гистограмм-

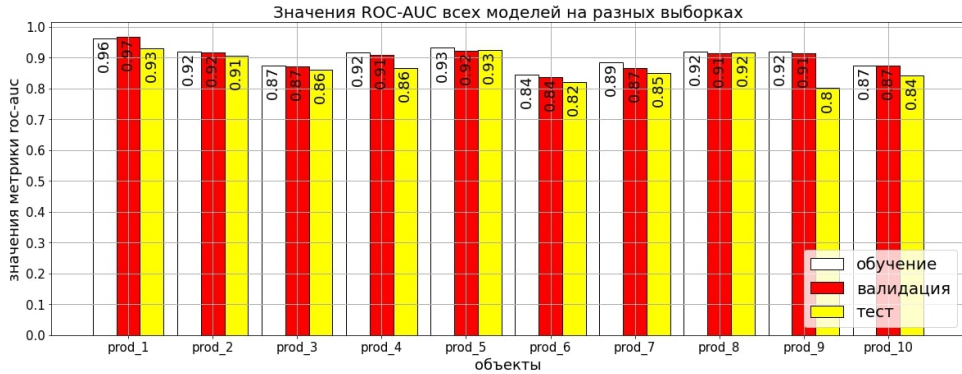


Рис. 1. Значения ROC-AUC всех моделей на разных выборках

ный метод [6, 7]. Меры качества итоговой рекомендательной системы: Hit Rate — 0,0859, Precision@3 — 0,0462 и Precision@5 — 0,0352.

Предложенный метод решает проблемы малого числа объектов при условии их взаимосвязанности, используя как коллаборативные данные, так и метаинформацию. Промежуточное качество классификаторов достаточное, а итоговая система показывает нелинейное повышение точности на первых k объектах с ростом k , что свидетельствует о правильном ранжировании.

Результат может быть использован для улучшения качества работы рекомендательных алгоритмов, работающих на подобных данных. Метод позволяет получить вероятности заинтересованности пользователей, которые могут быть также использованы для других целей. Метод гибок в плане добавления новых объектов и предоставляет широкие возможности для быстрого переобучения и повышения качества классификаторов.

- [1] *Al-bashiri H., Abdulhak M., Romli A., Hujainah F.* Collaborative Filtering Recommender System: Overview and Challenges // 3rd Int. Conf. Computational Intelligence & Communication Technology (CICT). — Ghaziabad, India, 2017. — Pp. 1–6.
- [2] *Zharova M. A., Tsurkov V. I.* Neural Network Approaches for Recommender System // J. Computer and Systems Sciences International. — Springer Nature, 2024. — Vol. 62. — Pp. 1048–1062.
- [3] *Ahn H., Kang H., Lee J.* Selecting a Small Number of Products for Effective User Profiling in Collaborative Filtering // Expert Systems with Applications. — Elsevier, 2010. — Vol. 37. — Pp. 3055–3062.
- [4] *Niculescu-Mizil A., Caruana R.* Predicting Good Probabilities with Supervised Learning // Machine Learning, Proc. 22nd International Conference (ICML). — Bonn, Germany, 2005. — Pp. 625–632.
- [5] *Dorogush A., Prokhorenkova L., Gusev G.* CatBoost: unbiased boosting with categorical features // arXiv. — 2019. — <https://arxiv.org/abs/1706.09516v5>

- [6] *Gupta C., Ramdas A.* Distribution-free calibration guarantees for histogram binning without sample splitting // arXiv. — 2021. — <https://arxiv.org/abs/2105.04656>
- [7] *Naeni M., Cooper G.* Classifier calibration: a survey on how to assess and improve predicted class probabilities // Machine Learning, Springer Nature, 2023. — T. 112, C. 3211–3260.

A Probabilistic Framework of Shared Decision-Making

*Pinhao Song*¹*

pinhao.song@kuleuven.be

¹Leuven, KU Leuven

This paper proposes a shared control framework for navigation with an omnidirectional agent. The key challenges of shared control are (i) distinguishing noise and the desired action, and (ii) acting when the shared controller disagrees with the user. To address the challenges, we model the shared policy and the expert user policy as stochastic distributions $\pi_S(\mathbf{a}|\mathbf{s})$ and $\pi_E(\mathbf{a}|\mathbf{s})$, where $\mathbf{a} \in \mathbb{R}^2$ is the velocity action.

Let us first define \mathbf{s} an abstract state that includes all the information needed to determine \mathbf{a} , and we will define it explicitly later, which does not affect the following deduction. We aim to minimize the following optimization goal:

$$\min_{\pi_S} \mathcal{D}_{KL}(\pi_S(\cdot|\mathbf{s}), \pi_E(\cdot|\mathbf{s})), \quad (1)$$

where $\mathcal{D}_{KL}(\cdot, \cdot)$ is KL divergence.

Unfortunately, analytically solving (1) is impossible because π_E cannot be explicitly obtained. Thus, we seek a data-driven way by maximizing log-likelihood instead:

$$\max \mathbb{E}_{(\mathbf{s}_{1:T+H}, \mathbf{a}_{1:T+H-1})} [\log p(\mathbf{s}_{T+1:T+H}, \mathbf{a}_{T:T+H-1} | \mathbf{s}_{1:T}, \mathbf{a}_{1:T-1})], \quad (2)$$

where \mathbf{s}_t is the state that includes the current dynamics and contextual information at the timestep t , and \mathbf{a}_t is the velocity action.

The expert trajectory $(\mathbf{s}_{1:T+H}, \mathbf{a}_{1:T+H-1})$ are sampled from the interaction between the user and the environment with failure samples filtered out, where $(\mathbf{s}_{1:T}, \mathbf{a}_{1:T-1})$ are the past trajectory and $(\mathbf{s}_{T+1:T+H}, \mathbf{a}_{T:T+H-1})$ are the future trajectory. Assuming the current expert user's intended action only depends on past T steps states, independent of past user's actions, we can decompose the likelihood in (2) as:

$$p(\mathbf{s}_{T+1:T+H}, \mathbf{a}_{T+1:T+H-1} | \mathbf{s}_{1:T}, \mathbf{a}_{1:T-1}) = \prod_{t=T}^{T+H-1} p(\mathbf{s}_{t+1} | \mathbf{s}_t, \mathbf{a}_t) p(\mathbf{a}_t | \mathbf{s}_{t-T+1:t}).$$

Introducing the latent state to represent the memory of states and a latent variable to represent multi-modality, we get:

$$\begin{aligned} p(\mathbf{a}_{T:T+H-1}, \mathbf{h}_{T:T+H} | \mathbf{s}_{1:T}) &= p(\mathbf{h}_T | \mathbf{s}_{1:T}) p(\mathbf{a}_{T:T+H-1}, \mathbf{h}_{T+1:T+H} | \mathbf{h}_T) = \\ &= p(\mathbf{h}_T | \mathbf{s}_{1:T}) \sum_{\mathbf{z}} p(\mathbf{z} | \mathbf{h}_{T+H}) \prod_{t=T}^{T+H-1} p(\mathbf{h}_{t+1} | \mathbf{h}_t, \mathbf{a}_t, \mathbf{z}) p(\mathbf{a}_t | \mathbf{h}_t, \mathbf{z}). \end{aligned}$$

Given the current user's command \mathbf{a}_T^u , we can infer the posterior distribution, as:

$$\begin{aligned} p(\mathbf{h}_{T+1:T+H}, \mathbf{a}_{T:T+H-1} | \mathbf{s}_{1:T}, \mathbf{a}_T^u) &\approx \\ &\approx \sum_{\mathbf{z}} p(\mathbf{h}_{T+1} | \mathbf{h}_T, \mathbf{a}_T, \mathbf{z}) p(\mathbf{a}_T | \mathbf{h}_T, \mathbf{z}, \mathbf{a}_T^u) q(\mathbf{z} | \mathbf{h}_T, \mathbf{a}_T^u) \cdot \\ &\quad \cdot \prod_{t=T+1}^{T+H-1} p(\mathbf{h}_{t+1} | \mathbf{h}_t, \mathbf{a}_t, \mathbf{z}) p(\mathbf{a}_t | \mathbf{h}_t, \mathbf{z}). \end{aligned}$$

Sampling \mathbf{a}_T from $p(\mathbf{h}_{T+1:T+H}, \mathbf{a}_{T:T+H-1} | \mathbf{s}_{1:T}, \mathbf{a}_T^u)$, we achieve a shared control of a user and an autonomous agent. The proposed framework considers the user's uncertainty to help filter out the noise, and the multi-modality of the proposed framework achieves high agreement with the user.

Новые методы построения рекомендательных систем на основе данных о покупках

*Шубочкин Евгений Иванович*¹

evgen.shubochkin@gmail.com

¹Москва, МГУ имени М. В. Ломоносова

В современных условиях розничной торговли, особенно в онлайн-сегменте, ключевым фактором успеха является способность персонализировать предложения для каждого клиента. В этом контексте рекомендательные системы играют важную роль, позволяя предлагать товары, которые наилучшим образом соответствуют предпочтениям покупателей. Одной из основных задач, связанных с рекомендациями, в индустрии является выбор рекомендаций в корзине. Как правило, в крупных онлайн магазинах цель рекомендательной полки в корзине — показать пользователю товары, которые больше всего подходят к текущему ее состоянию. Для таких рекомендаций логично использовать данные о чеках. В качестве бейзлайна для этой задачи используются подходы, основанные на частотности товаров, а также на ассоциативных правилах.

Метод построения рекомендательной модели, который был ранее предложен в [1], основывается на кластеризации товаров для получения их признакового

описания и градиентном бустинге. Для товара строится бинарный вектор, который соответствует чекам, в которых этот товар встретился. В векторном пространстве товаров проводится иерархическая кластеризация. В качестве признакового описания чека используются расстояния между товарами, содержащимися в чеке и кластерами. Для признакового описания товара-кандидата, используются расстояния между ним и кластерами.

Новый способ заключается в получении обучаемого векторного представления товаров, используя информацию о соседних товарах в чеке по аналогии с текстовой моделью Word2Vec [2], с той разницей, что в качестве контекста для обучаемого вектора товара используется информация о соседних товарах в чеке. Помимо признаков, полученных от кластеризации по аналогии с предыдущим методом, в новом способе используется расстояние между вектором чека, посчитанным как усредненный вектор товаров, содержащихся в нем, и вектором товара кандидата. Новый метод показывает перспективу подхода с получением обучаемых векторных представлений для товаров, и мотивирует дальнейшие исследования с применением более сложных нейросетевых архитектур.

В работе использован способ сборки датасета для обучения и оценки качества модели на основе чеков, предложенный в [1]. Предлагается случайно набирать товары-кандидаты из чека для позитивных примеров, причем исключая кандидата из чека, и случайно набирать товары-кандидаты для негативных примеров, из не содержащихся в чеке.

- [1] Zhuravlev Yu., Dokukin A., Senko O., Stefanovsky D., Saenko I., Korolev N. Comparison of Two Approaches to Recommender Systems with Anonymous Purchase Data // Supplementary Proceedings of the XXII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020). — CEUR-WS: 2020. — Vol. 2790. — Pp. 114–120.
- [2] Mikolov T., Chen K., Corrado G., Dean J. Distributed Representations of Words and Phrases and their Compositionality // Neural Information Processing Systems. — 2013. — Pp. 3111–3119.

Анализ графов целей и задач стратегического планирования

Шарафиев Артур Флюрович^{1*}

whiskeydudev@gmail.com

*Лемтюжникова Дарьяна Владимировна*¹

darabbbt@gmail.com

*Авдеева Зинаида Константиновна*¹

avdeeva@ipu.ru

¹Москва, ИПУ РАН

Из корпуса документов по стратегическому планированию экспертами извлечены и предложены для исследования цели и задачи, реализующие стратегию государственного развития РФ [1–3]. Цели и задачи формируют 9 графов, из которых 4 имеют и цели и задачи, а 5 — только цели.

В работе предлагается ряд характеристик для оценки качества стратегического плана по полученным графам: согласованность, сбалансированность и равнозначность. Под согласованностью двух целей понимается длина наибольшей последовательности задач, выполнение которых необходимо для достижения обеих целей, а под согласованностью графа — количество пар целей, согласованность которых больше нуля. Сбалансированность графа вычисляется как среднее нормированных разностей между количеством вершин в подряд идущих уровнях графа. Под равнозначностью понимается разность между числом целей и задач, необходимых для достижения некоторой цели, для целей с наибольшим и наименьшим числом таких целей и задач.

Проводится анализ взаимосвязи стандартных графовых индексов, как например, средняя степень вершины, коэффициент ассортативности, расстояние от графа до бинарного дерева, ранг матрицы смежности и т. д. с предложенными характеристиками. Для этого используется корреляционный и кластерный виды анализа. В рамках кластерного анализа используется алгоритм DBSCAN для построения кластеров графов в векторном пространстве. Элементами векторов данного пространства являются полученные значения стандартных индексов. Также проводится кластеризация графов отдельно по значениям каждого из стандартных индексов и каждой предложенной характеристики. Принимается, что между характеристикой и индексом существует возможная взаимосвязь, если абсолютное значение коэффициента корреляции между ними выше либо равно 0,85 и кластеризация по данному индексу соответствует кластеризации по характеристике с точностью до метки одного графа.

Из полученных результатов можно сделать вывод о возможных связях между индексами и характеристиками. Соответствующие связи представлены в таблице 1.

Таблица 1. Связи между индексами и характеристиками

Характеристика	Индекс
Равнозначность	Ранг матрицы смежности Средняя степень центральности Остовное число
Число согласованных пар целей	Средняя степень вершины Коэффициент кластерности Средняя относительная глубина сцепки
Сбалансированность	Низкая корреляция и несоответствие кластеров

В дальнейшем планируется проверить предложенные характеристики и провести аналогичный анализ с данными большего объема, взятыми, например, из [4].

[1] Указ Президента Российской Федерации от 02.07.2021 г. № 400. — <http://www.kremlin.ru/acts/bank/47046>

-
- [2] Указ Президента Российской Федерации от 08.11.2021 г. № 633. — <http://www.kremlin.ru/acts/bank/47244>
- [3] Федеральный закон от 28.06.2014 г. № 172-ФЗ. — <http://www.kremlin.ru/acts/bank/38630>
- [4] *Artemova E., Batura T., Golenkovskaya A., Ivanin V., Ivanov V., Sarkisyan V., Smurov I., Tutubalina E.* So What's the Plan? Mining Strategic Planning Documents // ArXiv. — 2020. — <https://arxiv.org/abs/2007.00257>

Анализ больших временных рядов с пропусками в реальном времени

*Семёнова Софья Эдуардовна*¹*

semenovasofyaeduardovna@gmail.com

*Голубцов Петр Викторович*¹

golubtsov@physics.msu.ru

*Шапкина Наталья Евгеньевна*¹

neshapkina@mail.ru

¹Москва, Московский государственный университет имени М. В. Ломоносова

При широкомасштабном сборе временных данных нередко приходится сталкиваться с пропущенными значениями. При этом большинство методов анализа временных рядов не рассчитаны на работу с рядами, имеющими пропуски. Поэтому довольно стандартной является предварительная обработка, состоящая в «заполнении» пропущенных значений [1]. К сожалению, такое заполнение обычно является эвристическим и может существенно исказить стохастические свойства временного ряда. В связи с этим становится актуальной разработка подходов и алгоритмов, которые могут полноценно работать с рядами, имеющими пропуски [2].

В современных исследованиях ситуация усугубляется еще и тем, что такие ряды все чаще имеют огромные объемы и непрерывно пополняются. Это переводит проблему в область потоков больших данных. Для эффективной обработки таких данных оказывается удобным разбить обработку данных на две фазы [3]. А именно, по мере поступления данных, в реальном времени, производить накопление некоторой информации, а окончательный результат обработки строить по мере необходимости на основании всей накопленной на данный момент информации.

В данной работе рассматривается алгоритм построения эмпирической автоковариационной функции стационарного временного ряда с пропусками. Предложенный подход использует идеи накопления информации в условиях больших данных [4]. А именно, путем анализа уже имеющихся данных временного ряда формируется специального вида каноническая информация, на основании которой легко вычисляется оценка автоковариационной функции. Далее, по мере поступления данных каноническая информация постоянно обновляется. В результате, актуальная оценка автоковариационной функции обновляется в режиме реального времени. При этом устраняется необходимость хранить в оперативном доступе полный набор данных. Общие принципы построения максимально эффективной формы представления информации, содержащейся в данных, рассмотрены в [5].

Построенная автоковариационная функция позволяет решать различные задачи. Например, на ее основании можно строить прогноз будущих значений или оптимальную оценку пропущенного фрагмента ряда по имеющимся соседним. В работе продемонстрирована работоспособность предложенного метода. Исследована зависимость точности построения автоковариационной функции от процента пропущенных значений. Наконец, рассмотрено влияние процента

пропусков на качество прогнозирования или заполнения пропущенных фрагментов временного ряда.

Отметим, что выделяемая каноническая информация обладает рядом замечательных свойств. В частности, она занимает компактный объем, не зависящий от объема представленных данных. Более того, фрагменты канонической информации можно «складывать». При этом, если два фрагмента канонической информации представляют два непересекающихся фрагмента временного ряда, то их сумма будет представлять объединенный временной ряд. Это позволяет накапливать информацию для большого ряда параллельно на нескольких компьютерах, разбив исходный ряд на фрагменты, а потом сложить полученные результаты и получить автоковариационную функцию для полного ряда.

- [1] *Грачев А. В.* К восстановлению пропусков в экспериментальных данных // Вестн. ННГУ им. Н. И. Лобачевского. Сер. Радиофизика. — 2004. — № 2. — С. 15–23.
- [2] *Дещеревский А. В., Журавлёв В. И., Никольский А. Н., Сидорин А. Я.* Проблемы анализа временных рядов с пропусками и методы их решения в программе WinABD // Геофизические процессы и биосфера. — 2016. — Т. 15, № 3. — С. 5–34.
- [3] *Голубцов П. В.* Понятие информации в контексте задач обработки больших данных // НТИ Сер. 2. Информационные процессы и системы. — 2017. — № 1. — С. 31–36.
- [4] *Golubtsov P.* Scalability and Parallelization of Sequential Processing: Big Data Demands and Information Algebras // Advances in Intelligent Systems and Computing. — Springer, 2020. — Vol. 1127. — Pp. 274–298.
- [5] *Голубцов П. В.* Оптимизация распределенной обработки больших данных: Алгебраические основы и понятие информации // Ученые записки физического факультета Московского Университета. — 2023. — № 5. — С. 2350101-1–2350101-7.

Модель автоматизированной системы имитации атак на основе машинного обучения

Греков Михаил Михайлович¹*

grekov.web@yandex.ru

Сычугов Алексей Алексеевич¹

xru2003@list.ru

¹Тула, Тульский государственный университет

Тестирование на проникновение (Penetration Testing) является важным аспектом сетевой безопасности, однако традиционные методы могут быть трудоемкими и дорогостоящими. Современные системы имитации взломов и атак (Breach and Attack Simulation, BAS) позволяют значительно автоматизировать данный процесс, однако большинство средств основаны на статических сценариях атак. Предлагается модель на основе генеративного состязательного имитационного обучения (Generative Adversarial Imitation Learning, GAIL) [1], а также усиление с помощью семантических вознаграждений и многоагентного подхода.

GAIL использует генератор и дискриминатор для создания реалистичных сценариев атаки [2]. Генератор обучается производить действия, которые неотличимы от экспертных, в то время как дискриминатор обучается правильно отличать действия эксперта от сгенерированных действий [3]. Одной из ключевых особенностей предложенной модели является способность генерировать реалистичные сценарии атак, которые могут имитировать поведение эксперта в области тестирования на проникновение. Это позволит снизить трудоемкость процесса и повысить его масштабируемость и адаптируемость.

Для улучшения модели вводятся семантические вознаграждения, которые оценивают успешность атак с точки зрения новизны и скрытности. Это позволяет агенту обучаться более сложным и реалистичным сценариям атак. Рассматривается возможность применения двух дискриминаторов и трансферного обучения для упрощения адаптации предобученной модели к новым сетевым инфраструктурам. Модель может использовать многоагентный подход, где несколько агентов действуют одновременно, что позволяет имитировать координированные атаки.

Используя предложенную модель BAS-системы, организации смогут получить более глубокое представление о защищенности информации и выявить потенциальные угрозы и уязвимости. Представленная модель предлагает ряд преимуществ, включая экономичность, масштабируемость и возможность генерировать реалистичные атаки.

- [1] *Ho J., Ermon S.* Generative Adversarial Imitation Learning // *Advances in Neural Information Processing Systems*. — NY: Curran Associates, 2016. — Vol. 29.
- [2] *Chen J.* GAIL-PT: A Generic Intelligent Penetration Testing Framework with Generative Adversarial Imitation Learning // *Computers & Security*. — UK: Elsevier Ltd, 2023. — Vol. 126 — Pp. 103055.
- [3] *Nguyen T.* Dual Discriminator Generative Adversarial Nets // *Advances in Neural Information Processing Systems*. — NY: Curran Associates, 2017. — Vol. 30.

Выделение эмбеддингов для активного обучения, соответствующих регионам интереса на изображении

Соколов Василий Александрович^{1,2*}

sokolov.va@phystech.edu

*Матвеев Иван Алексеевич*²

matveev@frccsc.ru

*Тепляков Лев Михайлович*²

teplyakov.l@ya.ru

¹Москва, МФТИ

²Москва, ФИЦ ИУ РАН

Современные нейронные сети для обработки изображений требуют больших наборов аннотированных данных. Для того, чтобы снизить затраты на разметку, применяются практики активного обучения — автоматического пополнения датасетов. Методология активного обучения включает в себя критерий выбора новых данных в размеченном наборе для последующей разметки и включения в обучающую выборку. Традиционные критерии отбора данных при работе с глубокими сетями для распознавания образов на изображениях встречаются такие трудности, как завышение уверенности модели, изменчивая структура признакового пространства, необходимость дообучать сети после добавления каждого изображения. Таким образом, актуальна проблема формулирования надежного критерия выбора данных для активного обучения для глубоких нейронных сетей.

Перспективным направлением решения этой задачи является работа с векторными репрезентациями (эмбеддингами) изображений, полученными из самих глубоких моделей. Такой подход, позволяющий выбирать новые изображения на основе расстояния в пространстве признаков, лишен части недостатков, присущих классическим критериям, т. к. не опирается на значение уверенности сети. Кроме того, метод позволяет отбирать новые данные не по отдельности, а группами, учитывая плотность их расположения в признаковом пространстве. При этом вектор признаков формируется для всего изображения, тогда как существует необходимость оценки ценности отдельных объектов на изображениях. Такие объекты могут быть выделены ограничивающими прямоугольниками либо масками сегментации. Предлагаемый в настоящей работе метод описывает способ получения эмбеддингов, соответствующих отдельным ограничивающим прямоугольникам на изображении.

Пирамидальная сеть выделения признаков (Feature Pyramid Network) [1] использует карты признаков, полученные на нескольких пространственных уровнях (масштабах). Для того, чтобы выделить на картах признаков регионы, соответствующие ограничивающим прямоугольникам на исходном изображении, используется метод понижения размерности карт активации в качестве альтернативы операции пулинга RoiAlign, предложенный в [2]. Т. к. этот метод использует билинейную интерполяцию вместо округления до ближайшей границы ячейки (грануляции), он также позволяет приводить регионы, полученные с разных уровней карт признаков, к одному и тому же размеру. Таким образом,

предложенное использование метода RoiAlign позволяет в рамках одной операции решить как задачу сопоставления карт активаций различного масштаба, так и уменьшить пространственные искажения. Комбинированные эмбединги понижаются в размерности с помощью метода PCA, после чего используются для подсчета расстояний в критерии отбора для активного обучения.

Работа поддержана грантом РФФ № 24-21-00330.

- [1] *Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S.* Feature Pyramid Networks for Object Detection // IEEE Conf. Computer Vision and Pattern Recognition (CVPR). — Honolulu: IEEE, 2017. — Pp. 936–944.
- [2] *He K., Gkioxari G., Dollár P., Girshick R.* Mask R-CNN // IEEE Int. Conf. Computer Vision (ICCV), Honkong: IEEE, 2017. — Pp. 2980–2988.

The Influence of Graph Construction Methods on the Generalization Ability of Graph Neural Networks in Neurophysiological EEG Analysis

Leonid Sidorov^{1*}

leon.sidorov@gmail.com

*Archil Maysuradze*¹

maysuradze@cs.msu.ru

¹Moscow, Lomonosov Moscow State University

In this work we study various methods for constructing an adjacency matrix for use in Graph Convolutional Networks (GCN). We validate different approaches in neurophysiology, specifically the analysis of electroencephalograms (EEG) to identify the P300 pattern. We compared “external” methods, which use information from outside the time series, with “internal” methods that rely solely on original EEG data.

Currently, “internal” methods, such as learning the adjacency matrix (graph learning, GL) [1, 2], are prevalent in EEG data processing with GCNs. In this method, the graph is constructed simultaneously with the training process. The adjacency matrix is typically parameterized using

$$A_{ij} = \frac{\exp(-ReLU(\mathbf{w}^T|\psi_i - \psi_j|))}{\sum_{k=1}^N \exp(-ReLU(\mathbf{w}^T|\psi_i - \psi_k|))},$$

where $\psi_i, \psi_j \in \mathbb{R}^T$ represent vector representations of the graph vertices, N is the total number of vertices, and $\mathbf{w} \in \mathbb{R}^T$ is a trainable weight vector. Regularization terms are added to the loss function:

$$\mathcal{L}_{gcn} = \alpha \underbrace{\frac{1}{2N^2} \sum_{i, j=1}^N \|\psi_i - \psi_j\|_2^2 A_{ij}}_{\text{locality}} + \beta \underbrace{\frac{(-1)}{N} \sum_{i=1}^N \log \left(\sum_{j=1}^N A_{ij} \right)}_{\text{connectivity}} + \gamma \underbrace{\frac{1}{2N^2} \|A\|_F^2}_{\text{sparsity}},$$

where $\alpha, \beta, \gamma \in \mathbb{R}$ are hyperparameters for locality, connectivity, and sparsity.

In EEG analysis, “external” methods often rely on the physical placement of electrodes. For example, in [3]

$$A_{ij} = \begin{cases} 1 & \text{if } A_{ij} \geq 1, \\ \frac{\delta}{d_{ij}^2} & \text{if } 0.1 \leq A_{ij} \leq 1, \\ 0.1 & \text{if } A_{ij} \leq 0.1, \end{cases} \quad (1)$$

where $d_{ij} \in \mathbb{R}$ is the Euclidean distance between vertices i and j , and $\delta \in \mathbb{R}$ is a connectivity factor. Also we incorporated the method applying *Delaunay triangulation* [4] to electrode coordinates, resulting in a planar graph consisting only of triangles. Each triangle’s circumscribed circle includes only its vertices, making the graph sparse.

Experiments show that a “hybrid” approach, where a graph obtained by “external” methods is adjusted during model training via gradient propagation (i. e. “internal” tuning), yields better results in recognizing the P300 pattern. “Hybrid” method accurately identifies the most informative channels, improving the generalization ability of GCNs. Combining Delaunay triangulation with this approach outperforms other initialization methods.

- [1] *Mathur P., Mittal T., Manocha D.* Dynamic Graph Modeling of Simultaneous EEG and Eye-Tracking Data for Reading Task Identification // ICASSP 2021-2021 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP). — Toronto: IEEE, 2021. — Pp. 1250–1254.
- [2] *Ye W., Zhang Z., Zhang M., Teng F., Zhang L., Li L., Huang G., Wang J., Ni D., Liang Z.* Semi-Supervised Dual-Stream Self-Attentive Adversarial Graph Contrastive Learning for Cross-Subject EEG-based Emotion Recognition // arXiv. — 2023. — <https://arxiv.org/abs/2308.11635>
- [3] *Jin M., Zhu E., Du C., He H., Li J.* PGCN: Pyramidal Graph Convolutional Network for EEG Emotion Recognition // arXiv. — 2023. — <https://arxiv.org/abs/2302.02520>
- [4] *Delaunay B. et al.* Sur la sphere vide // Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk. — M.: Izd. Akad. Nauk SSSR, 1934. — No. 6. — Pp. 793–800.

Метод маскировки входа для обучения моделей перевода

Скачков Николай Андреевич^{1*}

nikolaj-skachkov@yandex.ru

Воронцов Константин Вячеславович¹

vokov@forecsys.ru

¹Москва, ФИЦ ИУ РАН

Одной из важных задач в области анализа текстов на естественном языке является машинный перевод. Современные модели перевода основаны на нейросетях [1], обученных на миллионах пар переведенных предложений. Такие

модели обладают высоким качеством, но страдают систематическими ошибками при переводе [2].

В соответствии с [1], обучение модели машинного перевода можно рассматривать следующим образом:

$$\sum_{i=1}^N \sum_{j=1}^{|y_i|} \log P_{\theta}(y_i^j | y_i^{<j}, x_i) \longrightarrow \max_{\theta}, \quad (1)$$

где y_i — перевод текста x_i , $y_i^{<j}$ — его префикс до j -го слова, а $P_{\theta}(\cdot)$ — обучаемая модель перевода.

В задаче языкового моделирования метод обучения с маскированной функцией потерь дает лучшее качество на многих прикладных задачах, по сравнению с другими методами обучения [3]. Адаптируем идею маскированной функции потерь для задачи перевода:

$$L(X, Y) = \sum_{i=1}^N \sum_{j=1}^{|y_i|} \log P_{\theta}(y_i^j | y_i^{<j}, m(x_i), m(y_i)). \quad (2)$$

При обучении с функцией потерь (2) модель при генерации перевода видит входной и целевой тексты, в которых значительная часть последовательных токенов замаскирована. Маскирование входного текста заставляет модель внимательнее смотреть на те слова входа, которые остались доступны. Добавление маскированного перевода во вход модели позволяет учитывать правый контекст, что должно ограничивать зависимость генерации каждого нового слова от левого контекста. Более того, данный подход позволяет учитывать одноязычные данные в обучении.

Результат сравнения на корпусе WMT19 по метрике BLEU [4] моделей, обученных различными методами, можно увидеть в таблице. Можно заметить, что обучение модели перевода с маскированной функцией потерь (2) дало лучший результат по сравнению с обучением с функцией потерь (1), а добавление одноязычных данных на русском языке не принесло результата.

Модель	en-ru-wmt19, BLEU
Функция потерь (1)	34
Функция потерь (2)	34,35
Функция потерь (2) с одноязычными данными	33,5

Таким образом, в данной работе удалось показать, что обучение модели перевода с маскированной функцией потерь позволяет получить более высокое итоговое качество перевода. Предложенный подход также позволяет добавить одноязычные данные в обучение модели перевода, но роста качества таким способом добиться пока не удалось, вероятно, из-за большого количества параллельных данных на англо-русском направлении.

- [1] *Vaswani A., Shazeer N., Parmar N. et al.* Attention is All You Need // Proc. 31st Intern. Conf. Neural Information Processing Systems (NIPS'17). — Curran Associates Inc., Red Hook. N.Y, 2017. — Pp. 6000–6010.
- [2] *Yang Z., Cheng Y., Liu Y. et al.* Reducing Word Omission Errors in Neural Machine Translation: A Contrastive Learning Approach // Proc. 57th Annual Meeting of the Association for Computational Linguistics. — Florence, 2019. — P. 6191–6196.
- [3] *Raffel C., Shazeer N., Roberts A. et al.* Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer // J. Machine Learning Research. — Mountain View, 2020. — Vol. 21. — Pp. 1–67.
- [4] *Papineni K., Roukos S., Ward T., Zhu W.-J.* Bleu: a Method for Automatic Evaluation of Machine Translation // Proc. 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania. — USA: Association for Computational Linguistics, 2002. — Pp. 311–318.

Обнаружение объектов на изображениях с использованием сверточных сетей на графах и ограниченного количества тренировочных данных

Захаров Алексей Александрович^{1*}

aa-zaharov@ya.ru

*Захарова Мария Владимировна*¹

neurocv@ya.ru

*Жизняков Аркадий Львович*¹

lvovich1975@mail.ru

¹Муром, Муромский институт (филиал) федерального государственного бюджетного образовательного учреждения высшего образования «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

В настоящее время разработано много методов для обнаружения объектов на изображениях. Одними из наиболее известных методов обнаружения объектов на основе нейронных сетей являются YOLO, SSD, R-CNN и др. Однако возможности современных подходов сильно ухудшаются в реальных условиях наблюдения, а также при использовании ограниченного количества тренировочных данных. Часто необходимо обнаруживать объекты по категориям, у которых количество экземпляров в обучающем наборе ограничено. Подобные методы получили название *few-shot learning* (обучение с нескольких выстрелов).

Кроме того, многие существующие методы обнаружения объектов не используют информацию об окружающей сцене, которую можно задать с помощью графов [1]. Таким образом, в большинстве существующих методов обнаружения объектов отсутствуют знания о взаимосвязи категорий. Взаимосвязь категорий представляет собой информацию с нерегулярной структурой. Поскольку информация нерегулярно структурирована, знания о взаимосвязях могут быть хорошо обработаны сверточными сетями на графах (Graph Convolutional Network). Это позволит проанализировать взаимосвязи между объектами, что

увеличит робастность и точность разрабатываемых методов при использовании ограниченного количества тренировочных данных.

В разрабатываемом методе осуществляется построение графа знаний о взаимосвязях с помощью набора обучающих данных на основе условной вероятности между каждыми двумя категориями.

Разработанный метод обнаружения объектов включает два этапа: обнаружение объектов на основе моделей few-shot learning, анализ взаимосвязей между объектами сцены на основе сверточных сетей на графах для повышения точности обнаружения. При реализации разработанного метода с анализом взаимосвязей между объектами средняя точность (Mean Average Precision, mAP) обнаружения на тестовых изображениях увеличилась на 1–6 %.

Исследование выполнено за счет гранта Российского научного фонда № 23-21-00486, <https://rscf.ru/project/23-21-00486/>.

- [1] *Захаров А. А.* Метод сопоставления изображений с использованием тепловых ядер на графах // Автоматика и телемеханика. — 2022. — № 10. — С. 60–66.

Нейросетевая детекция линий разметки взлетно-посадочной полосы и рулежных дорожек летного поля

Зотов Андрей Владимирович^{1*}

ZotovAV@gosniias.ru

*Лебедев Максим Алексеевич*¹

MLebedev@gosniias.ru

*Комаров Денис Валерьевич*¹

DKomarov@gosniias.ru

¹Москва, ФАУ «ГосНИИАС»

В сфере авиации с каждым годом повышается требования к обеспечению безопасного движения воздушных судов (ВС) по рулежным дорожкам к месту стоянки или взлета. С увеличением количества воздушных перевозок и постоянно растущей загруженности аэропортов создается потребность в автоматизации движения ВС по аэродрому. Автоматизация руления воздушного судна по летному полю является комплексной задачей, состоящей из целого ряда подзадач, одна из которых — распознавание разметки взлетно-посадочной полосы (ВПП) и рулежных дорожек (РД). При этом работа алгоритма распознавания разметки ВПП и РД должна обеспечиваться при различных условиях видимости и быть устойчивой к визуальным (блики, засветы и т. д.) и механическим (износ, следы тормозного пути и т. д.) искажениям и дефектам.

Разметка осевой и боковых линий ВПП и РД схожа с разметкой полос движения автомобильных дорог, поэтому в данной работе предлагается применение перспективных нейросетевых алгоритмов определения полос движения к задаче распознавания разметки ВПП и РД. В работе рассматриваются нейросетевые алгоритмы [1–3], которые показали хорошие результаты на открытых датасетах определения полос движения автомобилей. Распознавание разметки на летном поле и рулежных дорожках в системах технического зрения так же может быть

применено для следующих задач: уточнение локализации ВС на основе визуальной информации, помощь пилотам при движении в условиях ограниченной видимости, оптимизация и контроль маршрутов движения на территории аэродрома.

Данная исследовательская работа показывает перспективность использования нейросетевых методов для решения задач, связанных с управлением ВС и регулированием их движения на территории аэродрома, что может значительно повысить безопасность и эффективность работы аэропорта.

- [1] *Tu Zheng, Yifei Huang, Yang Liu, Wenjian Tang, Zheng Yang, Deng Cai, Xiaofei He* CLRNet: Cross Layer Refinement Network for Lane Detection // CVPR. — 2022.
- [2] *Honda H., Uchida Y.* CLRerNet: improving confidence of lane detection with LaneIoU // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. — 2024. — Pp. 1176–1185.
- [3] *Jin D., Park W., Jeong S.-G., Kwon H., Kim C.-S.* Eigenlanes: Data-Driven Lane Descriptors for Structurally Diverse Lanes // in Proc. IEE CVPR. — 2022.

Нейросетевая детекция взлетно-посадочной полосы на изображениях от системы технического зрения воздушных судов гражданской авиации

*Воробьев Тимур Константинович*¹*

VorobievTK@gosniias.ru

*Ермаков Павел Григорьевич*¹

ErmakovPG@gosniias.ru

*Фурс Егор Васильевич*¹

Furs_Egor@gosniias.ru

¹Москва, ФАУ «ГосНИИАС»

На сегодняшний день бортовые системы технического зрения (СТЗ) можно разделить на 4 поколения. Примерно с 2016 года начал формироваться облик интеллектуальной СТЗ 4 поколения, в основе которого лежит применение современных высокочувствительных датчиков разных спектральных диапазонов и разной физической природы, высокопроизводительных вычислителей с наличием современного графического ускорителя и функционального программного обеспечения, обеспечивающего интеллектуальную информационную поддержку экипажа воздушного судна (ВС) на основе нового поколения алгоритмов обработки визуальной информации закабинного пространства.

Применение глубоких нейронных сетей в интеллектуальных СТЗ позволяет решать широкий круг задач, в том числе, например, осуществлять автоматическое обнаружение взлетно-посадочной полосы (ВПП) с помощью компьютерного зрения, что в свою очередь позволяет обеспечивать коррекцию данных глобальной навигационной спутниковой системы (ГНСС) на основе заранее известных параметров ВПП. Это позволяет фактически осуществлять полностью автоматическую посадку с навигацией и визуальным контролем без использования наземных систем.

В данной работе представлен нейросетевой метод обнаружения ВПП на этапе захода на посадку и посадки. Обнаружение ВПП осуществляется в два этапа: нейросетевой поиск области с ВПП на изображении от СТЗ на базе архитектуры семейства YOLO [1] с последующим нейросетевым уточнением угловых точек образа ВПП с использованием архитектуры MnasNet [2]. Отличительной особенностью предлагаемого подхода является наличие дополнительного анализа обнаруженных элементов с последующей фильтрацией для получения уточненного геометрического положения ВПП, а именно положения 4 угловых точек.

- [1] Wang C. Y., Bochkovskiy A., Liao H. Y. M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. — 2023. — Pp. 7464–7475.
- [2] Tan M. et al. Mnasnet: Platform-Aware Neural Architecture Search for Mobile // Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition. — 2019. — Pp. 2820–2828.

Метод объединения специализированных нейросетей в множество многозадачных моделей путем выявления взаимосвязей между их внутренними представлениями

Панов Алексей Сергеевич^{1*}

alpan@bk.ru

Матвеев Иван Алексеевич¹

ivanmatveev@mail.ru

¹Москва, ФИЦ ИУ РАН

Многозадачное обучение — область машинного обучения, объединяющая методы построения и обучения многозадачных моделей, моделей одновременно решающих сразу несколько различных задач. Преимуществами многозадачных моделей по сравнению с однозадачными моделями являются более эффективное использование памяти и вычислительных ресурсов, а также более эффективное использование данных, т. к. данные для одной задачи могут содержать полезную информацию для другой. В связи с этим основной особенностью многозадачного обучения является необходимость учета взаимосвязей между задачами. Существующие на текущий момент методы поиска задач, которые могут эффективно решаться в рамках одной многозадачной модели, требуют большого количества вычислительных ресурсов [1].

В данной работе предложен эффективный метод установления взаимосвязей между задачами путем анализа внутренних представлений однозадачных нейросетей. В качестве меры близости внутренних представлений нейросетей используется индекс СКА (Centered Kernel Alignment) [2]. Показано, что агрегирующее значение индексов СКА между соответствующими слоями двух однозадачных нейросетей можно использовать как меру близости этих нейросетей и соответствующих им задач. Определив попарные близости для некоторого множества задач, можно выполнить кластеризацию на данном множестве таким образом, что задачи внутри каждого кластера будут эффективно решать-

ся в рамках одной многозадачной модели. Преимуществом данного метода является отсутствие необходимости дополнительного обучения моделей на этапе анализа, и, как следствие, вычислительная эффективность. При этом качество работы предложенного метода не уступает качеству наиболее известных и эффективных методов, например [1].

Также предложен метод, который на основе анализа внутренних представлений нейросетей, осуществляет поиск наиболее эффективного объединения однозадачных нейросетей в одну многозадачную нейросеть. В данном случае предполагается, что объединяемые сети имеют одинаковую общую часть — бэкбон — состоящую из некоторого количества блоков (слоев). А задача сводится к оптимальному выбору блоков, которые будут общими для двух или более нейросетей. При этом оптимальность определяется как наилучшее значение агрегированной по всем задачам метрики качества. В предложенном методе используется эмпирически полученная производная от индекса СКА величина — SCKA. Определив значение SCKA для каждой пары соответствующих блоков двух нейросетей, можно установить наилучшее разбиение на общие и индивидуальные блоки нейросетей. Далее данный подход обобщается на произвольное количество нейросетей. Показано, что качество данного метода сравнимо с качеством полного перебора всех возможных вариантов, но при этом вычислительная эффективность метода существенно выше, т. к. не требуется обучение каждого из возможных вариантов.

- [1] *Fifty C., Amid E., Zhao Z., Yu T. et al.* Efficiently Identifying Task Groupings for Multi-Task Learning // 35th Conf. Neural Information Processing Systems (NeurIPS 2021). — Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2022.
- [2] *Kornblith S., Norouzi M., Lee H., Hinton G.* Similarity of Neural Network Representations Revisited // Proc. 36th Int. Conf. Machine Learning. — Long Beach, California: Proceedings of Machine Learning Research, 2019.

Обучение устойчивой к геометрическим искажением входных данных нейросетевой модели классификации изображений

Дмитриенко Анна Евгеньевна^{1*}

dmitrienko.ae@phystech.edu

*Гнеушев Александр Николаевич*¹

algneushev@gmail.com

¹Москва, ФИЦ ИУ РАН

Геометрические искажения изображений объектов существенно снижают эффективность моделей машинного обучения, предназначенных для анализа и классификации изображений. В данной работе исследуется проблема неустойчивости моделей к поворотам и предлагается решение на основе нейросетевых архитектур. Для решения этой проблемы часто используют дообучение на аугментированных данных, что зачастую способствует переобучению и вле-

чет уменьшение точности на оригинальных данных. Основная цель исследования — разработать модель, способную эффективно классифицировать изображения, включая те, что повернуты на небольшой угол.

В работе предлагается использование обучаемого трансформационного блока (Spatial Transformer Network, STN) для модификации пространственного распределения внутренних признаков нейросетевой модели, основанной на архитектуре визуального трансформера (ViT). Более конкретно, архитектура модели изменилась только в последнем блоке, а именно, была распараллелена на две ветви: первая ветвь принимает оригинальные изображения только во время обучения, а вторая ветвь с STN блоком — повернутые на случайный угол до 20° . Обучение модели происходит таким образом, чтобы итоговые признаки и выходы ветвей были идентичны между собой. Итоговая формула для функции потерь выглядит следующим образом:

$$\mathcal{L} = \frac{1}{M} \sum_{j=1}^M w_j \text{CE}(y, \hat{y}_j) + \frac{1}{M-1} \left(\sum_{i=2}^M w_{f_i} \text{MSE}(f_1, f_i) + w_{l_i} \text{MSE}(l_1, l_i) \right),$$

где M — количество ветвей модели (в данной работе $M = 2$), l_i — выходы i -й ветви модели, f_i — признаки последнего блока энкодера i -й ветви модели, w_j — веса для классификационной функции потерь i -й ветви, w_{f_i} , w_{l_i} — веса для функций потерь признаков и логитов i -й ветви, y — оригинальные метки класса, CE — классификационная функция потерь, отвечающая за основную задачу распознавания объектов, независимо от их ориентации, MSE — среднеквадратичная функция потерь, специально нацеленная на обучение инвариантности к вращению. Таким образом, совместная оптимизация этих двух функций потерь позволяет обучить модель, устойчивую к вращению объектов.

Эксперименты показывают, что новая модель обладает значительно улучшенной устойчивостью к поворотам до 20° : точность классификации снижается всего на $0,2\%$ (при исходном значении 3%).

Cluster Merging Techniques for the Travelling Salesman Problem

*Egor Barashov*¹

barashov.eb@gmail.com

¹Moscow, V. A. Trapeznikov Institute of Control Science RAS

The traveling salesman problem (TSP) is a classic optimization problem that has been extensively studied in the literature. The objective is to find the shortest possible route that allows the salesman to visit all given cities and return to the starting point.

When the problem becomes large and includes additional constraints, such as clustering of cities, traditional methods may become inefficient. In such cases, clustering is used to simplify the problem: cities are divided into clusters, and a sub-problem TSP is solved within each cluster. Then, the challenge is to combine these solutions into a single solution for the entire problem. This aspect has received little attention in the literature, which is the main focus of this work.

In this study, we explore various cluster merging procedures based on known algorithms [1] and propose our own technique, which is improved from Sigal's work [2] by reducing the number of cluster iterations.

The merging procedure consists of two stages:

1. Selection of a pair of clusters for merging: based on the proximity of the centroids of the clusters (centroids are recalculated after each merge); based on the proximity of the clusters to the newly formed cluster. In the first case, clusters are merged simply based on proximity, while in the second one, expansion proceeds from the point of the first merge.

2. Merging procedure. Various algorithms are used: based on the proximity of the middle edges of the clusters for merging, based on the proximity of the edges to the centroid of the cluster with which the merge is occurring, or by reducing the total path by removing two edges within two clusters and adding two new edges that form the merge.

To conduct the experiments, TSP instances from the open TSPLib library were used. These instances were divided into clusters using two methods: k-means and affinity propagation. Then, a route was built within each cluster using the open solver Or-tools, and the merging procedure began.

The results show that our developed algorithm significantly reduces the total distance compared to the methods on cluster merging based on the proximity of edges to the centroid and based on the proximity of the centers of edges of two clusters. On a dataset with 144 vertices *pr144.tsp*, an average relative improvement of 8.89% was obtained, for a dataset with 439 vertices *pr439.tsp*, an average improvement of 3.57% was obtained. For datasets with fewer vertices, an average improvement of at least 2.83% was obtained. This confirms the effectiveness of our approach compared to traditional methods, demonstrating improvements in solution accuracy.

Thus, the proposed methodology and new cluster merging algorithms for the TSP are useful for solving large and complex problems, enhancing route optimization and logistics.

The study was carried out with partial financial support of the RSF under the scientific project 22-71-10131.

- [1] *Anaya Fuentes G. E. et al.* Solution to Travelling Salesman Problem by Clusters and a Modified Multi-Restart Iterated Local Search Metaheuristic // PloS one. — 2018. — Vol. 13, No. 8. — Pp. e0201868.
- [2] *Sigal I. H.* Combined Algorithms for Solving the Traveling Salesman Problem // Moscow: All-Russian Center of the USSR Academy of Sciences. — 1985.

Обобщение полиномиального случая выпуклой оболочки для задачи коммивояжера: метод попарных сравнений

Красоткин Семён Александрович¹★

krasotkin.sa@phystech.edu

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

Задача коммивояжера является задачей дискретной оптимизации, в которой требуется определить кратчайший маршрут, проходящий через все заданные вершины (города), посещая каждую только один раз. В общем случае эта задача относится к NP-трудным и не имеет полиномиального алгоритма для ее решения. Однако существуют специальные случаи, для которых можно найти оптимальное решение за полиномиальное время.

Для расширения применимости алгоритмов и изучения структуры NP-трудных задач в статье [1] предложен *метод попарного сравнения*. Этот метод основывается на сравнении двух примеров задачи A и B по определенному критерию с использованием *функции различия*. Значение этой функции соответствует значению целевой функции задачи A при применении полиномиального (или псевдополиномиального) точного решения задачи B . Таким образом, можно оценить, насколько эвристика, обеспечивающая точное решение для подкласса задач, применима к более широкому классу задач. В работе рассматривается применение метода попарного сравнения к задаче коммивояжера.

Анализируется специальный случай задачи, когда все вершины расположены на евклидовой плоскости и образуют границу выпуклого многоугольника [2], то есть лежат на выпуклой оболочке множества вершин. В этом случае оптимальный маршрут проходит по границе многоугольника, и решение можно найти за полиномиальное время $O(n)$, где n — количество вершин.

Предлагается обобщить этот специальный случай на задачу коммивояжера в классической постановке путем нахождения такого евклидового пространства, в котором выполняется неравенство треугольника, а вершины располагаются на границе политопы в этом пространстве.

Для перехода от исходной задачи к специальному случаю был исследован подход, заключающийся в рассмотрении исходной задачи как набора вложенных политопов.

Для данного варианта перехода предложена функция попарного сравнения и описан алгоритм получения допустимого решения.

Работа поддержана грантом РФФИ № 22-71-10131.

- [1] *Lemtyuzhnikova D., Chebotarev P., Goubko M., Shushko N., Somov M.* Pairwise Similarity Estimation for Discrete Optimization Problems // *Advances in Systems Science and Applications.* — 2023. — Vol. 23, No. 2. — Pp. 164–177.
- [2] *Burkard R. E. et al.* Well-Solvable Special Cases of the Traveling Salesman Problem: A Survey // *SIAM review.* — 1998. — Vol. 40, No. 3. — Pp. 496–546.

Метод попарного сравнения для задачи двухуровневого голосования

Шушко Никита Игоревич^{1*}

shushko.ni@phystech.edu

¹Москва, Институт проблем управления им. В. А. Трапезникова РАН

В социальных и технических системах для принятия решений нередко используются процедуры двухуровневого голосования. В таких процедурах на первом этапе голосование проводится в локальных группах агентов, на втором этапе результаты голосования групп агрегируются — также посредством голосования — в итоговое решение [1]. Один из центральных вопросов при анализе двухуровневых процедур: «При какой минимальной доле агентов, поддерживающих предложение, оно может быть в итоге принято данной процедурой?» Разность числа агентов, поддерживающих предложение, и агентов, не поддерживающих предложение, называют числом доминирования.

Рассматривается граф $G = (V, E)$, $|V| = n$, каждая вершина которого имеет петлю. Вершина $v \in V$ имеет множество соседей $N_v = \{w \in V : (v, w) \in E\}$, в которое входит и v . Каждая вершина $v \in V$ обладает мнением $f(v) \in \{-1, 1\}$.

Функция мнений расширяется на подмножества $W \subseteq V$ как сумма мнений вершин множества W по формуле $f(W) = \sum_{v \in W} f(v)$. В ходе голосования вершина $v \in V$ учитывает мнение вершин из N_v , и голосует «за», если $f(N_v) > 0$, либо «против» в ином случае. Если число $|V^+| = |\{v \in V : f(N_v) \geq 1\}|$, голосующих «за», строго больше $|V|/2$, предложение, поставленное на голосование, принимается. Числом доминирования называется величина $\gamma_{\text{maj}}(G) = \min\{f(V) : |V^+| > |V|/2\}$.

Рассматривается задача поиска такой функции мнений $f(v)$ графа G , которая реализует число доминирования $\gamma_{\text{maj}}(G)$, т. е. обеспечивает успех голосования при минимальном числе вершин v таких, что $f(v) = 1$.

В [2] предлагается эвристический алгоритм решения задачи поиска числа доминирования с гарантированной абсолютной погрешностью, основанный

на методе попарного сравнения. В данной работе предлагается уточнить лемму об удалении ребра из графа, которая была сформулирована в [2], и предложить модификацию алгоритма с более точной функцией попарного сравнения, что позволяет сделать более точную оценку абсолютной погрешности алгоритма.

Лемма 1 Пусть G — непустой граф, а H — граф, получаемый из G удалением произвольного ребра uv . Тогда

— если $f(u) = f(v) = 1$, то

$$-4 \leq \gamma(G) - \gamma(H) \leq 0;$$

— если $f(u) = f(v) = -1$, то

$$0 \leq \gamma(G) - \gamma(H) \leq 2;$$

— в остальных случаях

$$-2 \leq \gamma(G) - \gamma(H) \leq 2.$$

Работа поддержана грантом РФФИ № 22-71-10131.

- [1] *Chebotarev P., Peleg D.* The Power of Small Coalitions under Two-Tier Majority on Regular Graphs // arXiv. — 2022. — <https://arxiv.org/abs/2210.03410>
- [2] *Lemtyuzhnikova D., Chebotarev P., Goubko M., Shushko N., Somov M.* Pairwise Similarity Estimation for Discrete Optimization Problems // Advances in Systems Science and Applications. — 2023. — Vol. 23, No. 2. — Pp. 164–177.

О «сжатых» степенных рядах и сложности их вычисления

*Карацуба Екатерина Анатольевна*¹

ekaratsuba@gmail.com

¹Москва, Федеральный исследовательский центр «Информатика и управление»
Российской академии наук

Под «сжатыми» степенными рядами, будем иметь в виду ряды вида

$$W_r(x) = \sum_{j=0}^{\infty} d(j)x^{j^r}, \quad |x| < 1, \quad r = 2, 3, \dots \quad (1)$$

Такие ряды возникают в разных областях математики и физики, например, в квантовой оптике (см. [1, 2]). В теории чисел и теории специальных функций такими рядами являются: эйлерова тэта-функция и тэта-ряды Якоби (см., например, [3]), а также функция $\omega(x)$, введенная Риманом (см., например, [4]),

$$\omega(x) = \sum_{j=1}^{\infty} e^{-j^2 \pi x},$$

с помощью которой дзета-функцию Римана можно приблизить в рациональных точках.

В настоящей работе представлены алгоритмы и оценки сложности вычисления степенных рядов вида (1), в которых участвуют лишь члены с переменной в степени, равной степени натурального числа, не меньшей 2; для аргумента x разной арифметической природы. Под алгоритмом ниже предполагается правило или способ вычисления без формализации этого понятия.

Посредством двух совершенно разных подходов к аппроксимации «сжатых» степенных рядов получены алгоритмы с оценками сложности вычисления одинакового вида. Доказана [5]

Теорема. Сложность вычисления ряда

$$\widetilde{W}_r(x) = \sum_{j=0}^{\infty} x^{j^r}, \quad |x| < 1, \quad r = 2, 3, \dots,$$

где x — трансцендентное число, есть

$$s_{\widetilde{W}_r(x)}(n) = O\left(n^{1+\frac{1}{r}+\varepsilon}\right),$$

для любого $\varepsilon > 0$ и $n > n_1(\varepsilon)$, $n \rightarrow \infty$.

Поскольку в теории сложности вычислений до сих пор (июль 2024) не было найдено ни одной оценки сложности вычисления снизу (кроме тривиальных), трудно утверждать, что быстрого алгоритма вычисления «сжатого» степенного ряда с трансцендентным аргументом не существует. Однако, то, что оценка

сложности вычисления такого ряда посредством двух совершенно разных методов оказалась одинаковой по n , $n \rightarrow \infty$, с точностью до констант, стоящих в O , свидетельствует о том, что полученная оценка близка к «естественной» (возможно даже к оптимальной) для вычисления «сжатого» степенного ряда с трансцендентным аргументом.

- [1] *Dodonov V. V.* “Nonclassical” states in quantum optics: a “squeezed” review of the first 75 years // *J. Opt. B Quantum Semiclass. Opt.* — 2002. — Vol. ,No. 1. — R1–R33.
- [2] *Tana R.* “Nonclassical states of light propagating in Kerr media, Chapter 6”, *Theory of Nonclassical States of Light*, eds. Dodonov V. V., Man’ko V. I. — CRC Press, 2019. — Pp. 267–312.
- [3] *Ватсон Дж. Н., Уиттекер Э. Т.* Курс современного анализа. — М.: Мир, 1963.
- [4] *Воронин С. М., Карацуба А. А.* Дзета-функция Римана. — М.: Наука, 1994.
- [5] *Карацуба Е. А.* О сложности вычисления «сжатых» степенных рядов // *Математические Заметки.* — М.: Изд. МИРАН им. В. А. Стеклова, 2023. — Т. 114, № 1. — С. 113–120.

Оценивание качества сегментации цифровых изображений

Мурашов Дмитрий Михайлович^{1*}

d_murashov@mail.ru

Мурашова Елизавета Дмитриевна¹

murlizon@yandex.ru

¹Москва, ФИЦ ИУ РАН

Для анализа качества сегментации цифровых изображений широко применяются базы тестовых изображений, которые включают также эталонные разбиения, полученные экспертами. Качество оценивается с помощью таких показателей, как мера Жаккара, коэффициент Шимкевича-Симпсона, индекс Рэнда, вариация информации, и др. [1]. Наряду с вычислением показателей качества разбиений, порождаемых различными алгоритмами сегментации, важно оценить теоретически достижимые граничные значения этих показателей для анализируемых изображений при известном эталонном разбиении на сегменты, формируемом экспертами вручную. Такие граничные значения могут быть использованы для сравнения разных эталонов и выбора наилучших из них для настройки и анализа качества алгоритмов сегментации. В представляемой работе предлагается использовать методологию, описанную в работе [2] для получения соотношения «взаимная информация — вероятность ошибки сегментации» и оценки теоретически достижимой минимальной ошибки сегментации по информационным характеристикам сегментируемого изображения и его эталонного разбиения.

Пусть имеется изображение U из N пикселей, его эталонное (Ground Truth) разбиение V^{GT} на K^{GT} сегментов и разбиение V на K сегментов, полученное некоторым алгоритмом сегментации.

Предполагается, что некоторая сцена может быть представлена идеальным (эталонным) изображением V^{GT} . Пусть с помощью измерительной системы получено зашумленное изображение U . Алгоритм сегментации формирует из изображения U разбиение V . Предполагается, что V^{GT} , U и V — дискретные случайные переменные. Эта модель может быть представлена следующей марковской цепью: $V^{GT} \rightarrow U \rightarrow V$. Требуется оценить вероятность ошибки сегментации в виде зависимости от информационных характеристик изображений V^{GT} и U . Вероятность ошибки сегментации по мере Хемминга зависит от количества пикселей, ошибочно размеченных при формировании сегментов S_j^V разбиения V относительно сегментов $S_i^{V^{GT}}$ эталонного разбиения V^{GT} .

Получена нижняя граница функции $J(P_e)$, задающей соотношение между вероятностью ошибки сегментации P_e и взаимной информацией $I(V^{GT}; U)$, аналогичная границам функций «скорость — погрешность» в задаче кодирования и «взаимная информация — погрешность» в задаче классификации. В работе также получены оценки границ величины вариации информации. В частности, найдена зависимость максимального значения вариации информации между разбиениями V^{GT} и V от вероятности ошибки P_e . Вычислена оценка минимального значения вариации информации между разбиениями V^{GT} и V , равная

условной энтропии. Проведен вычислительный эксперимент с использованием тестовых изображений и их эталонных разбиений из базы BSDS500, подтвердивший справедливость полученных границ. Полученные в работе теоретические границы позволяют оценить достижимые значения показателей качества сегментации по информационным характеристикам анализируемых изображений и эталонных разбиений, а также сравнить различные эталонные разбиения.

- [1] Zhang H. Image Segmentation Evaluation: A Survey of Unsupervised Methods // Computer Vision and Image Understanding. — Elsevier Inc., 2008. — Vol. 110, No. 2. — Pp. 260–280.
- [2] Ланге А. М., Ланге М. М., Парамонов С. В. О соотношении взаимной информации и вероятности ошибки в задаче классификации данных // Ж. вычисл. матем. и матем. физ. — М: Наука, 2021. — Т. 61, № 7. — С. 1192–1205.

Оптимизация декомпозиции изображений по топологическим признакам с использованием персистентных ландшафтов

Абакумов Артём Владимирович^{1*}

artem210966@yandex.ru

Еремеев Сергей Владимирович¹

sv-eremeev@yandex.ru

¹Муром, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

Декомпозиция по топологическим признакам позволяет при малых вычислительных ресурсах сегментировать изображения на основе соотношений яркостей соседних пикселей [1]. Результатом декомпозиции будет набор компонент — структур, содержащих топологические свойства (период существования) и матрицу пикселей, которые принадлежат компоненте. Компоненты строятся иерархически в виде дерева. Их можно использовать для разных задач, в том числе для классификации изображений [2].

Один из недостатков топологической декомпозиции — это большой объем компонент, особенно шумовых. Также зачастую один объект выделяется несколькими компонентами, что мешает обработке. Подобные проблемы можно решать фильтрацией, однако, подбирать параметры отсеивания приходится индивидуально для каждого случая.

В данной работе предлагается метод автоматического уменьшения числа компонент без искажения топологических свойств путем использования персистентных ландшафтов [3]. Персистентный ландшафт — это математическая модель, основанная на переводе вышеупомянутых компонент в гильбертово пространство по формуле $\lambda_k(t) = v_{kt}$, где k — индекс ландшафта, v_t — упорядоченное по убыванию множество высот $\{h_{t1}, h_{t2}, \dots, h_{ti}, \dots\}$, $h_{ti} \geq 0$ — высота,

равная $\min(t - b_i, d_i - t)$, $[b_i, d_i]$ — период существования i -й компоненты ($i = 1, 2, \dots, N$), t — значение яркости ($t = 0, 1, \dots, 255$), для которого определяется высота ландшафта, N — количество компонент.

Данные вычисления порождают кривые, как на рис. 1, которые можно рассматривать как компоненты ландшафтов, но в них отсутствует информация о матрицах. Так как каждый элемент матрицы — это яркость, то можно расположить эти яркости на шкале периода существования компонент. Учитывая, что при формировании ландшафтов берутся определенные промежутки у периода существования компонент, можно также включать в ландшафт все элементы из матрицы, чья яркость принадлежит этому диапазону. В результате многие компоненты объединяются, за счет чего и достигается оптимизация. Экспери-

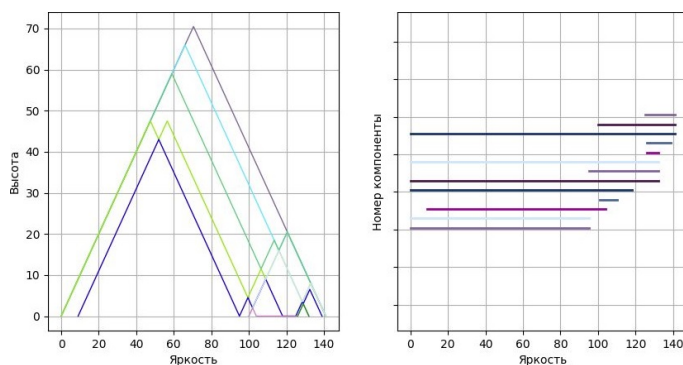


Рис. 1. Персистентные ландшафты и исходные компоненты

менты показали, что количество компонент сокращается с 47% до 72% без потери топологической информации. Тестирование проводилось на изображениях спутниковых снимков и фотографиях разных объектов. Таким образом, после предложенной предобработки можно не только вычислить ландшафты, но и матрицы, которые помогут получить более качественную декомпозицию.

Исследование выполнено за счет гранта Российского научного фонда № 23-21-10064, <https://rscf.ru/project/23-21-10064/>.

- [1] Еремеев С. В., Абакумов А. В., Андрианов Д. Е., Титов Д. В. Метод разложения изображения по топологическим признакам // Компьютерная оптика. — 2022. — Т. 46, № 6. — С. 939–947. <https://doi.org/10.18287/2412-6179-CO-1080>.
- [2] Еремеев С. В., Абакумов А. В. Классификация объектов на изображениях с учетом искажений на основе двухэтапного топологического анализа // Научно-технический вестник информационных технологий, механики и оптики. — 2022. — Т. 22, № 1. — С. 82–92. <https://doi.org/10.17586/2226-1494-2022-22-1-82-92>.
- [3] Bubenik P., Dlotko P. A Persistence Landscapes Toolbox for Topological Statistics // J. Symbolic Computation. — 2017. — Vol. 78. — P. 91–114. <https://doi.org/10.1016/j.jsc.2016.03.009>.

Применение топологической декомпозиции изображений для классификации криогелей поливинилового спирта

Панкратов Денис Андреевич^{1*}
*Еремеев Сергей Владимирович*¹

denis_pankratov2000@mail.ru

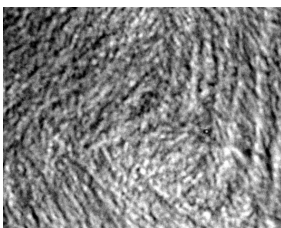
sv-eremeev@yandex.ru

¹Муром, Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых»

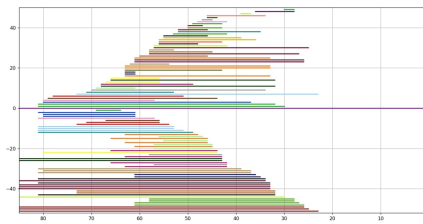
На сегодняшний день одной из основных задач в исследовании криогелей поливинилового спирта (рис. 1а) является разработка алгоритмов контроля их качества на основе анализа снимков полимеров. Данный анализ осложнен тем, что при различных способах получения криогелей их текстура может сильно отличаться при схожих физико-химических свойствах [1, 2]. Для более точного анализа снимков необходимо провести их классификацию.

В качестве способа получения исходных признаков для классификации используется метод яркостного разложения изображения по топологическим особенностям на отдельные компоненты связности, совокупность которых представляет собой баркод [3, 4]. Основная суть метода лежит в просмотре пикселей изображения по мере убывания яркости. При нахождении данного значения яркости на изображении создается компонента. При соприкосновении нескольких компонент происходит поглощение одной компоненты другой.

Для увеличения точности алгоритма в построении баркода помимо поглощенных компонент используются также поглощаемые. Далее полученные компоненты кластеризуются при помощи метода k-средних. Результат кластеризации представлен на рис. 1б, где на оси Ox отражается яркость пикселей, Oy — номер компоненты, а каждому классу компонент соответствует свой цвет.



(а) Изображение криогеля



(б) Классификация компонент баркода криогеля

Рис. 1. Криогели поливинилового спирта

Для классификации криогелей использовались такие характеристики, как количество классов компонент, количество классов с поглощенными и поглощаемыми компонентами, а также данные количества, где компонент больше одной, количество поглощенных компонент с разным количеством пикселей, и количество поглощаемых компонент с различным временем существования.

В качестве алгоритма классификации в данной работе использовался метод опорных векторов. Обучение проводилось на 20 классах. Объем обучающей выборки составил 1000 изображений, тестовой — 200. В ходе экспериментов удалось добиться точности распознавания 75 %.

Исследование выполнено за счет гранта Российского научного фонда № 23-21-10064, <https://rscf.ru/project/23-21-10064/>.

- [1] Лозинский В. И., Дамшквал Л. Г., Шаскольский Б. Л. и др. Изучение криоструктурирования полимерных систем. 27. физико-химические свойства криогелей поливинилового спирта и особенности их макропористой морфологии // Коллоидный журнал. — 2007. — Т. 69, № 6. — С. 798–816.
- [2] Лозинский В. И., Дамшквал Л. Г., Курочкин И. Н., Курочкин И. И. Изучение криоструктурирования полимерных систем. 28. физико-химические свойства и морфология криогелей поливинилового спирта, сформированных многократным замораживанием-оттаиванием // Коллоидный журнал. — 2008. — Т. 70, № 2. — С. 212–222.
- [3] Еремеев С. В., Андрианов Д. Е., Титов Д. В. Алгоритм совмещения пространственных объектов разномасштабных карт на основе топологического анализа данных // Компьютерная оптика. — 2019. — Т. 43, № 6. — С. 1021–1029. <https://doi.org/10.18287/2412-6179-2019-43-6-1021-1029>
- [4] Еремеев С. В., Абакумов А. В., Андрианов Д. Е., Титов Д. В. Метод разложения изображения по топологическим признакам // Компьютерная оптика. — 2022. — Т. 46, № 6. — С. 939–947. <https://doi.org/10.18287/2412-6179-CO-1080>

Интегрированный метод сегментации и уточнения рельефа атомных изображений на основе свертки формы зонда и нейронной сети архитектуры U-net

Зырянов Константин Сергеевич¹* zyryanov.konstantin.sergeevich@gmail.com

Орешкин Андрей Иванович¹ oreshkin@spmlab.phys.msu.su

Чуличков Алексей Иванович¹ achulichkov@gmail.com

¹Москва, МГУ имени М. В. Ломоносова

Одной из основных задач в области обработки изображений атомного масштаба является точная сегментация и определение формы атомов на поверхности [1]. В данной работе предложен новый подход к обработке изображений сканирующей зондовой микроскопии, состоящий в комбинации методов восстановления рельефа поверхности и сегментации изображений. Разработан алгоритм, осуществляющий наиболее точное восстановление рельефа поверхности, путем аппроксимации формы зонда параболоидом с параметрами, доставляющими наивысшую точность восстановления рельефа.

Для этого моделируется поверхность с известным рельефом, затем производится численное сканирование зондом с формой в виде параболоида. Далее

выполняется свертка отсканированной поверхности с инвертированной формой зонда. Результатом такой численной деконволюции является более точный скан поверхности [2]. Выбираются такие параметры формы зонда, при которых разница между истинной формой рельефа и восстановленной является минимальной.

Для обучения нейронной сети был создан специализированный набор данных, состоящий из 500 изображений атомов, ориентированный на задачу сегментации. С использованием платформы RoboFlow была выполнена разметка этого набора данных, в результате чего получены маски изображений атомов, при этом размер каждого изображения составили 640×640 пикселей. Данные были разделены на обучающую, тестовую и валидационную выборки в соотношении 7 : 2 : 1. Перед обучением были применены техники аугментации данных, включая размытие по Гауссу, зеркальное отображение, вращение изображения, вырезание части изображения и изменение яркости. Исследованы различные размеры нейросети с архитектурой U-net для сегментации изображений [3].

Классический метод основан на использовании адаптивного порогового значения и морфологической операции закрытия. Данный метод представляет собой эффективную стратегию для обработки и выделения объектов на изображениях. Применяется адаптивный порог, который делит изображение на участки и вычисляет пороговое значение для каждого участка индивидуально, учитывая локальные интенсивности пикселей. Для улучшения сегментации и устранения мелких артефактов, таких как черные дыры внутри объектов белого цвета или разрывы между объектами, применяется морфологическая операция закрытия.

Преимуществом сети U-net является многослойный перцептрон, который используется для получения представления признаков изображения на краю объекта, что позволяет обновить начальное обнаружение, в результате получаются более тонкие края объектов.

При обучении были заданы следующие гиперпараметры: размер пакета — 8, начальный шаг обучения — 0,0003 и количество эпох — 25.

Метод	mAP, %	Время, с
Нейронная сеть U-net	68,7	1,946
Классический метод	48,3	0,004

Производительность моделей оценивается с использованием показателя mAP. mAP рассчитывается как средняя точность модели при различных пороговых значениях IoU (пересечение над объединением). В таблице показаны результаты валидации двух методов для сегментации атомов на изображениях сканирующей зондовой микроскопии. Значение показателя mAP для сегментации после обучения нейросети архитектуры U-net на 20 % выше, чем у классического метода. Результаты экспериментов показывают, что нейросеть с более высокой точностью решает задачу сегментации атомов, однако имеет более низкую скорость обработки изображений.

- [1] *Binnig G., Rohrer H.* Scanning Tunneling Microscopy // *Helv. Phys. Acta.* — 1982. — No. 55. — Pp. 726–735.
- [2] *Williams P. M., Shakesheff K. M. et al.* Blind Reconstruction of Scanning Probe Image Data // *J. Vac. Sci. Technol.* — 1996. — B 14. — Pp. 1557–1562.
- [3] *Ronneberger O., Fischer P., Brox T.* U-Net: Convolutional Networks for Biomedical Image Segmentation // In: Navab N., Hornegger J., Wells W., Frangi A. (eds) *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015. Lecture Notes in Computer Science.* — Springer, Cham, 2015. — Vol. 9351.

Генерация оптимальных приближений серого изображения

Харинов Михаил Вячеславович¹

khar@iias.spb.su

¹Санкт-Петербург, СПб ФИЦ РАН

Для инвариантного относительно изменения масштаба полииерархического представления объектов, определяемого независимо от способа своего вычисления, решается NP-трудная задача генерации *оптимальных* приближений в $g = 1, 2, \dots$ тонах с минимальными значениями E_g суммарной квадратичной ошибки E . При этом также развивается общее решение [1] для цветных изображений, т.к. алгоритмы обработки в терминах слияния и разделения надвое множеств пикселей сохраняют смысл для цветных изображений.

В случае серого (полутонного) изображения решение выражается последовательностью разбиений яркостной шкалы на непрерывные диапазоны (рис. 1). На рис. 1 по горизонтальной оси отсчитываются ненулевые бины гистограммы,



Рис. 1. Стандартное изображение (слева) и квадратная матрица оптимальных разбиений на диапазоны (справа), в строках которой правые граничные бины гистограммы каждого диапазона помечены черным цветом.

а по вертикальной оси сверху вниз — значения числа яркостных диапазонов в разбиении шкалы яркости, задаваемом соответствующей строкой матрицы. Примечательно, что на рис. 1 проявляется образ гистограммы изображения.

В программе реализации для вычисления оптимизированного приближения как функции любого исходного в соревновательном режиме применяются алгоритмы снижения E_g при неизменном g . Они обеспечивают генерацию по результирующему приближению *структурированной* иерархии приближений, которая его содержит и описывается монотонной выпуклой последовательностью ошибок E_g . Множество структурированных иерархий, ошибки E_g которых мажорируют ошибки целевой неиерархической последовательности оптимальных приближений, что обеспечивает ее эффективное вычисление.

Исследование выполнено при поддержке государственного бюджета, проект № FFZF-2022-0006.

- [1] Харинов М. В. Генерация оптимальных приближений цветного изображения // Всерос. конф. ММРО-21, Москва, 12–15 декабря 2023. — С. 115–117 (расширенная версия доклада в печати).

О задаче восстановления функциональной зависимости при наличии интенсивной импульсной помехи

Снитковский Максим Владиславович^{1*}

snitkovskymaks@gmail.com

*Чулчихов Алексей Иванович*¹

achulichkov@gmail.com

*Фаломкин Игорь Игоревич*²

igor.falomkin@gmail.com

¹Москва, МГУ им. М. В. Ломоносова

²Москва, ООО «Ай Ти Ви групп»

В различных областях методы компьютерного зрения применяются для определения траекторий движущихся объектов. Траекторию требуется определить по набору измерений положения объектов в разные моменты времени. При этом положение измеряется с погрешностью и отсутствует информация, к какому объекту относится каждое измеренное положение. Подобные задачи возникают в робототехнике, видеонаблюдении, при интерпретации данных физических экспериментов. В докладе рассматриваются математические методы и алгоритмы решения этой задачи.

В работе будем считать, что модель погрешности измерения положения объекта представляет собой сумму широкополосного шума и импульсов, амплитуда которых существенно больше дисперсии шума. Второе слагаемое будем называть выбросами. Целью исследования является анализ эффективности методов, выявление методов, наиболее устойчивых к шуму и выбросам в данных. В работе математические методы применяются к решению задачи определения траекторий движения объектов в поле зрения неподвижной видеокамеры. В этом

случае для анализа передается последовательность кадров. На каждом кадре определяются положения объектов, по определенным положениям строятся траектории движения.

Для нахождения траекторий одиночных объектов можно использовать метод наименьших квадратов (МНК) [1]. Этот метод минимизирует сумму квадратов отклонений измеренных положений от искомой траектории. Однако МНК чувствителен к выбросам в данных, что может приводить к значительным ошибкам в определении траекторий. Для решения этой проблемы в работе применяется алгоритм RANSAC, который устойчив к выбросам благодаря итеративному процессу выбора подмножества данных из начального набора измерений, снижающему влияние выбросов на результат аппроксимации. Кроме того, МНК не применим для случая, когда в последовательности кадров присутствует более чем один объект, в то время как RANSAC можно обобщить на случай одновременного движения в поле зрения видеокамеры нескольких объектов [2].

Еще одним мощным инструментом для выявления геометрических фигур (в том числе и траекторий движения), таких как прямые и окружности, в данных с выбросами и шумом является преобразование Хафа [3, 4]. Основная идея метода заключается в преобразовании исходных данных, представляющих собой положения объектов в разные моменты времени, в параметрическое пространство, где геометрические фигуры представляются в виде точек, принадлежащих этому параметрическому пространству. Метод позволяет эффективно определять траектории объектов, несмотря на наличие шума в данных. Преобразование Хафа также позволяет выявлять множество геометрических фигур одновременно, что делает его полезным при анализе сложных сцен с несколькими объектами. Однако оно также имеет свои недостатки, связанные с квантованием пространства параметров, а также с многомерностью параметрических пространств для большинства классов кривых, в которых ищутся аппроксимирующие траектории, например, полиномам n -й степени соответствует $(n + 1)$ -мерное пространство. Это приводит к неэффективности метода, когда требуется высокая точность, в условиях плотного расположения объектов.

В заключение предлагается ряд подходов к улучшению рассмотренных методов. К таким улучшениям относятся учет типа анализируемых данных, использование априорной информации о характере движения объектов и др.

- [1] *Mutun I. B., Rusafov B. C.* Анализ и обработка экспериментальных данных: учебно-методическое пособие. — М.: Физический факультет МГУ, 2002. — 44 с.
- [2] *Kolmogorov V., Zabih R.* What Energy Functions Can Be Minimized via Graph Cuts? // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2004. — Pp. 5–25.
- [3] *Illingworth J., Kittler J.* Hough Transform for Robust Regression and Automated Detection // CVGIP. — 1988. — Pp. 5–25.

- [4] *Beniamini O., Burshtein D.* Fast Hough Transform and Approximation Properties of Dyadic Patterns // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2017. — Pp. 5–25.

Improving Text Detection by Predicting Missed Bounding Boxes Using Graph Models

Artyom S. Matveev^{1,2}

matveev.as@phystech.edu

Ivan A. Matveev^{2*}

matveev@frccsc.ru

¹Dolgoprudny, Moscow Institute of Physics and Technology

²Moscow, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences

An actual problem in the text detection field is either an inaccurate detection of a text region or a complete miss of a whole bounding box. The latter issue can be explicitly noticed in complex or low-quality images, leading to the loss of information and incomplete text recognition results. In this paper, we propose a novel approach to enhance text detection by predicting missed bounding boxes using a graph-based technique. We represent all the detected text boxes as nodes of a graph. The edges of a graph connect nodes of the same nature. For example, if our data are documents, then the detected bounding boxes are split into a range of node sets, each of which represents a specific type of observed data in an image (e. g., personal data and field names in a payment receipt). Further, we estimate the detection results, and if we find that some bounding boxes are missing (e. g., by incorporating a priori knowledge received after the text recognition step), then we consider such a graph to be incomplete. Incomplete graphs are tackled by a graph completion model [1] that predicts the likely positions of missed text boxes in an image. By enriching the initial text detection results with these additional bounding boxes, our approach reduces the likelihood of missing important textual information, leading to more reliable outcomes on various downstream tasks, such as text recognition, and others.

We conduct extensive experiments on open access text detection datasets and evaluate the reliability of the proposed technique on a number of text detection models such as TextBPN++ [2], DBNet++ [3], CRAFT [4], and a few others.

This method offers a promising direction for improving text detection systems, particularly in challenging scenarios where some text regions may be missed.

- [1] *Tran C., Shin W.-Y.* Network Completion: Beyond Matrix Completion // 2021 Int. Conf. Information Networking (ICOIN). — Jeju Island, Korea (South): IEEE, 2021. — Pp. 667–670.
- [2] *Zhang S.-X., Yang C., et al.* Arbitrary Shape Text Detection via Boundary Transformer // IEEE Transactions on Multimedia. — United States: IEEE, 2023. — Vol. 26. — Pp. 1747–1760.

- [3] *Liao M., Zou Z., et al.* Real-Time Scene Text Detection With Differentiable Binarization and Adaptive Scale Fusion // IEEE Transactions on Pattern Analysis and Machine Intelligence. — United States: IEEE, 2023. — Vol. 45. — Pp. 919–931.
- [4] *Baek Y., Lee B., et al.* Character Region Awareness for Text Detection // 2019 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR). — Long Beach, CA, United States: IEEE, 2019. — Pp. 9357–9366.

Генерация текстового описания ОКТ-снимков на русском языке с использованием визуально-языковых моделей

*Волков Егор Николаевич*¹*

envolkoff@gmail.com

*Аверкин Алексей Николаевич*¹

averkin2003@inbox.ru

¹Москва, ФИЦ ИУ РАН

Использование технологий искусственного интеллекта (ИИ) в медицине переживает небывалый рост. Продукты на основе ИИ уже внедрены в анализ медицинских данных различной модальности (текст, таблицы, изображения, видео, звук). В последние годы задача автоматической генерации описания медицинских изображений с использованием визуально-языковых моделей (от англ. Vision-Language Models) становится одной из наиболее актуальных в части внедрения технологий ИИ в медицине [1]. Уже разработаны и активно применяются в клинической практике подобные решения для снимков лучевой диагностики.

Тем не менее, остаются некоторые области медицины, где процесс внедрения подобных решений затруднен по многим причинам. Одной из таких областей является офтальмология. Несмотря на существование большого количества исследований, подтверждающих преимущества нейросетей в автоматическом анализе офтальмологических изображений (фундус-, ОКТ-снимки) [2], найдены лишь единичные публикации по разработке решений для автоматической генерации описания таких изображений, что говорит о необходимости проведения исследований в данном направлении.

В данном исследовании в качестве базовой визуально-языковой модели была использована архитектура VLP, благодаря ее широким возможностям к адаптации под разные целевые области [3]. Также были исследованы возможности составных энкодер-декодер моделей с учетом их применения к данным на русском языке. Оценка эффективности работы моделей производилась на основе расчета показателей качества машинного перевода BLEU и ROUGE. Обучение моделей производилось на созданном специально для данной задачи датасете пар «изображение» (ОКТ-снимок) — «текст» (описание ОКТ-снимка на русском языке). Необходимость создания собственного датасета была продиктовано отсутствием в открытом доступе подобных наборов данных для любого языкового домена.

В ходе проведенного исследования были показаны перспективы использования визуально-языковых моделей в офтальмологии, в частности, впервые применительно к задаче автоматической генерации описания ОКТ-снимков на русском языке. Приводится процесс формирования набора данных, выбора архитектуры моделей, их обучения. Обсуждаются возможности и перспективы дальнейшего улучшения результатов работы моделей.

Работа поддержана грантом РФФИ № 24-21-00330.

- [1] *Hartsock I., Rasool G.* Vision-Language Models for Medical Report Generation and Visual Question Answering: A Review // arXiv. — 2024. — <https://arxiv.org/abs/2403.02469>
- [2] *Averkin A. N., Volkov E. N., Yarushev S. A.* Explainable Artificial Intelligence in Deep Learning Neural Nets-Based Digital Images Analysis // J. Comput. Syst. Sci. Int. — Springer, 2024. — Pp. 175–203.
- [3] *Fields C., Kennington C.* BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation // arXiv. — 2022. — <https://arxiv.org/abs/2201.12086>

Формирование интерпретируемых признаков на основе минимального остовного леса и их применение для определения структурированности расположения наночастиц на снимках с электронного микроскопа

Курбаков Михаил Юрьевич^{1*}

muwsik@mail.ru

*Сулимова Валентина Вячеславовна*¹

vsulimova@yandex.ru

*Середин Олег Сергеевич*¹

oseredin@yandex.ru

*Копылов Андрей Валериевич*¹

and.kopylov@gmail.com

¹Тула, Тульский государственный университет

Одной из важных практических задач химической промышленности является выявление скрытых дефектов на поверхности углеродных материалов, влияющих на их качество. Технология решения данной задачи заключается в нанесении на поверхность материала металлических наночастиц с последующим анализом их взаимного расположения, зафиксированного на изображениях, полученных со сканирующего электронного микроскопа (СЭМ-изображениях) [1]. Согласно исследованиям [1] наличие структурированности во взаимном расположении наночастиц свидетельствует о наличии скрытых дефектов, которые не могут быть обнаружены другими методами.

В литературе известны попытки решения задачи определения структурированности напрямую, используя нейросетевые методы различной степени сложности [2]. Однако нейросетевые модели имеют очень большое число параметров, результаты их работы плохо интерпретируются, а полученные модели могут

опираться на области изображения, не содержащие наночастицы, что вызывает недоверие к таким методам со стороны химиков-исследователей.

В рамках предыдущих исследований нами был предложен метод детектирования наночастиц на СЭМ-изображении [3], а также подход к формализации признаков структурированности, основанный на идентификации линий, образуемых детектированными наночастицами [4].

В данной работе, более детально описанной в [5], мы усовершенствуем предложенный в [4] подход, дополняя его новыми интерпретируемыми признаками, получаемыми на основе минимального остовного леса.

В результате предложенный подход включает в себя расчет достаточно простых признаков структурированности из трех основных групп: ориентация наночастиц, построение ломаных линий с использованием кратчайшего незамкнутого пути и минимального остовного леса. В дальнейшем эти признаки используются для определения факта структурированности с использованием метода SVM.

Экспериментальное исследование проводилось на базе СЭМ-изображений [6] для нескольких различных материалов, содержащих 750 изображений со структурированным и 250 изображений с неструктурированным расположением наночастиц. В наборе данных имеются изображения разного масштаба с различной степенью плотности расположения наночастиц.

В результате анализа факта структурированности были достигнуты высокие показатели качества классификации изображений, заметно превосходящие результаты, полученные при помощи нейронных сетей [2]. Более того, было продемонстрировано, что для достижения высокой точности классификации достаточно лишь небольшого количества предложенных признаков структурированности.

Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания FEWG-2024-0001. Авторы благодарят научную школу академика В. П. Ананикова за тему исследования, полезные дискуссии и экспериментальные данные.

- [1] *Pentsak E. O., Kashin A. S., Polynski M. V., Kvashnina K. O., Glatzel P., Ananikov V. P.* Spatial Imaging of Carbon Reactivity Centers in Pd/C Catalytic Systems // *Chem. Sci.* — 2015. — Vol. 6. — Pp. 3302–3313.
- [2] *Boiko D. A., Pentsak E. O., Cherepanova V. A., Gordeev E. G., Ananikov V. P.* Deep Neural Network Analysis of Nanoparticle Ordering to Identify Defects in Layered Carbon Materials // *Chem. Sci.* — 2021. — No. 12. — Pp. 7428–7441.
- [3] *Kopylov A. V., Kurbakov M. Y., Seredin O. S., Sulimova V. V., Boko D. A., Cherepanova V. A., Pentsak E. O., Ananikov V. P.* Automated Recognition of Nanoparticles in SEM images of Pd/C Catalysts // *Nanomaterials.* — 2022. — Vol. 12, No. 21. — P. 3914.
- [4] *Kurbakov M. Y., Sulimova V. V., Kopylov A. V., Seredin O. S., Boiko D. A., Pentsak E. O., Cherepanova V. A., Ananikov V. P.* Determining the Orderliness

of Carbon Materials with Nanoparticle Imaging and Explainable Machine Learning // *Nanoscale*. — 2024. — Vol. 16. — Pp. 13663–13676.

- [5] *Kurbakov M. Y., Sulimova V. V., Kopylov A. V., Seredin O. S.* Graph Method for Interpretable Determining Nanoparticles Ordering in Electron Microscopy Images // *Computer Optics* (in print).
- [6] *Boiko D. A., Pentsak E. O., Cherepanova V. A., Ananikov V. P.* Electron Microscopy Dataset for the Recognition of Nanoscale Ordering Effects and Location of Nanoparticles // *Sci Data*. — 2020. — Vol. 8, No. 101.

Алгоритмы морфологического анализа векторизованных границ на изображениях

Рижтер Андрей Александрович^{1*}

urfin17@yandex.ru

Мурынин Александр Борисович^{1,2}

AMurynin@bk.ru

¹Москва, Научно-исследовательский институт аэрокосмического мониторинга «АЭРОКОСМОС»

²Москва, Федеральный исследовательский центр «Информатика и управление» РАН

Выделение границ объектов является важной задачей анализа изображений, решение которой позволяет получать информацию о физических характеристиках исследуемых объектов [1]. Проблема в том, что не существует единого подхода к автоматической векторизации изображений, и большая часть обработки данных до сих пор производится вручную.

Введем понятие путь, как упорядоченный набор всех точек границы объекта, в котором соседние точки восьмисвязны или четырехсвязны. В настоящей работе предлагается метод и алгоритмы анализа векторной информации, базирующиеся на морфологических операциях, выполняемых над путями.

Предлагается комплекс алгоритмов, реализующих векторные морфологические операции, разделенные на следующие блоки:

- простые операции: инверсия, ротация, замыкание, удлинение и др.;
- по распознаванию путевых структур: ключевых точек, повторных точек, петель, отростков, узловых связок и др.;
- по деградации пути: выпрямление (линеаризация), фигурация, нормализация; удаление путевых структур;
- над участками пути: сложение и вычитание, умножение и деление путей, сегментация пути;
- по изменению траектории пути: преобразования движения (сдвиги, повороты), растяжения, подобия (симметрии), поляризации (ортогонализации);
- квазибинарные операции: изменение связности, передискретизация, дилатация и эрозия, открытие / закрытие, скелетизация и др.;
- групповые операции: сегментация, паттернизация, агрегация, фильтрация, интерполяция и экстраполяция, конъектура и др.

Алгоритмы реализованы в исследовательском макете программного обеспечения, которое позволяет рассчитывать и анализировать векторную информацию разного типа на изображениях.

Приведены примеры работы алгоритмов при обработке контурных изображений. Так, операции по деградации пути, имеющие характер «косметических», корректирующих операций, удаляют неинформативные структуры пути и приводят векторную информацию к виду, лучшему для визуального восприятия. В частности, выпрямление — операция, выполняемая по каждому сегменту пути в отдельности, состоящая в последовательном удалении точек в составе сегмента, имеющих наименьшее отклонение от отрезков из двух соседних с ними точек.

В качестве примера данный метод был применен к векторизации бумажной картографической информации, контурных изображений, нарисованных от руки.

Работа выполнена при поддержке АО «Тазмар АйТи-солюшнз», г. Санкт-Петербург.

- [1] Воробьев В. Е., Мурынин А. Б., Рихтер А. А. Выбор методов кластеризации при машинном обучении для исследования экологических объектов по спутниковым данным // Известия РАН. Теория и системы управления. — 2024. — № 5.

Декомпозиция сложных акустических атмосферных сигналов на части определенной формы

Закиров Марат Нафисович^{1*}

zakirov.mn16@physics.msu.ru

Чуличков Алексей Иванович^{1,2}

achulichkov@gmail.com

Куличков Сергей Николаевич^{1,2}

snik1953@gmail.com

¹Москва, Институт физики атмосферы имени А. М. Обухова РАН

²Москва, Московский государственный университет имени М. В. Ломоносова

Инфразвуковые сигналы генерируются в результате многих природных и антропогенных процессов. Благодаря малому затуханию низкочастотная составляющая акустических сигналов может распространяться в слоях атмосферы на многие сотни километров и нести информацию как о мощных процессах, происходящих на Земле, таких как взрывы, извержения вулканов, землетрясения, так и о структуре атмосферы, от разнородных слоев которой отражается сигнал.

Обнаружение инфразвуковых сигналов на фоне шума является довольно сложной задачей [1]. Ее сложность связана с наличием значительного фонового шума, а также с тем, что волны от импульсных источников приходят в приемник по разным траекториям, при этом претерпевая существенные амплитудные искажения.

В настоящей работе предлагается математическая модель регистрируемого инфразвукового сигнала, в которой он является суммой шумовой составляющей и линейной комбинации N - и U -волн известной формы, сдвинутых друг относительно друга на некоторые временные интервалы. Задача состоит в оценке количества сдвинутых N - и U -волн, коэффициентов линейной комбинации и времен задержек. Шум предполагается белым гауссовым с заданной дисперсией.

Метод решения состоит из двух этапов. На первом этапе считается, что заданы профили N - и U -волн, так, что регистрируемый сигнал представим в виде:

$$\xi(t) = \sum_{k=1}^K [a_k N_k(t - t_{0,k}) + b_k U_k(t - t_{0,k})] + \nu(t), \quad t = t_1, \dots, t_n$$

или в матричной форме: $\xi = Af + \nu$, где $f = \{a_1, b_1, \dots, a_K, b_K\}$. Далее методами теории измерительно-вычислительных систем [2] строится наилучшая оценка вектора f (т. е. коэффициентов линейной комбинации). Эта оценка имеет вид $\hat{f} = R\xi$, где линейное преобразование R является решением задачи на минимум:

$$\inf_{R \in L} \{ \|RA - I\|_2^2 \mid \|R\nu\|^2 \leq \varepsilon \},$$

в которой минимизируется отличие RA от тождественного преобразования I при ограничении на величину среднего квадрата нормы шума $R\nu$. $\|RA - I\|_2^2$ — квадрат нормы Гильберта-Шмидта линейного оператора $RA - I$ [2].

На втором этапе выбираются координаты вектора $\hat{f} = R\xi$, модуль которых превосходит заданный порог. Все координаты, модуль которых ниже порога, обрезаются в ноль. В результате такой фильтрации мы получаем наиболее значимые компоненты N- и U- инфразвукового сигнала.

В докладе приводятся примеры декомпозиции реальных сигналов в виде линейной комбинации нескольких N- и U-волн, демонстрирующие эффективность предложенного метода. Знание времен прихода сигналов N- и U-волн позволяет оценить структуру слоистой горизонтально однородной атмосферы.

Работа выполнена по госзаданию, тема 17.1.

- [1] Чучузов И. П., Куличков С. Н. Распространение инфразвуковых волн в анизотропной флуктуирующей атмосфере. — М.: ГЕОС, 2020. — 260 с.
- [2] Пытьев Ю. П. Методы математического моделирования измерительно-вычислительных систем. — М.: Изд-во Московского университета, 2024. — 431 с.

Сравнение алгоритмов поиска аномалий, основанных на методах машинного обучения и спектрального анализа

*Караваяева Ксения Эдуардовна*¹

karavaeva.ke21@physics.msu.ru

*Макеев Даниил Александрович*¹*

makeev.da21@physics.msu.ru

*Шапкина Наталья Евгеньевна*¹

shapkina@physics.msu.ru

*Голубцов Пётр Викторович*¹

golubtsov@physics.msu.ru

¹Москва, МГУ им. М. В. Ломоносова

Современные методы исследования позволяют получать данные с высоким разрешением по времени. Однако в данных различным образом могут появляться аномалии. Для временных рядов можно ввести классификацию аномалий на три типа [1]: одиночные, групповые и условные (контекстуальные). В этой работе рассматриваются методы выявления одиночных (точечных) аномалий.

В работе исследуются алгоритмы, способные обрабатывать неразмеченные данные. В качестве исходных данных выступают синтетические временные ряды с известными параметрами: суперпозиция гармоник или сглаженный белый шум.

Для сравнения алгоритмов используются следующие метрики: recall (полнота), precision (точность) и F1-score [2]. Recall демонстрирует способность алгоритма обнаруживать класс аномалий, а precision — способность отличать аномалии от обычных данных. Метрика F1-score представляет собой агрегированный критерий качества модели.

Тестирование производительности алгоритмов осуществляется в среде Python 3 Google Compute Engine backend (System RAM — 12,7 GB).

Задачу поиска аномалий можно рассматривать как задачу прогнозирования с дальнейшим сравнением исходных и спрогнозированных данных. Для постро-

ения прогноза используется модель экспоненциального сглаживания и линейная регрессия.

Также задачу поиска аномалий можно рассматривать как задачу кластеризации данных. Данные, лежащие вне кластеров, можно считать аномальными. В качестве примера такого подхода используется алгоритм KNN [3].

Еще одним вариантом решения задачи о нахождении единичных выбросов является использование частотного анализа. В работе, в качестве такого метода рассматривается метод сингулярного анализа спектра (SSA) [4], который легко обобщается на случай многомерных временных рядов.

Результаты сравнения скорости и эффективности работы рассмотренных алгоритмов представлены в таблице 1. В результате исследования можно отметить, что алгоритм SSA является наиболее точным из рассмотренных. В свою очередь метод прогнозирования позволяет выявлять аномальные значения в режиме реального времени, а алгоритм KNN является самым быстрыми из рассмотренных, и может использоваться для первичных оценок.

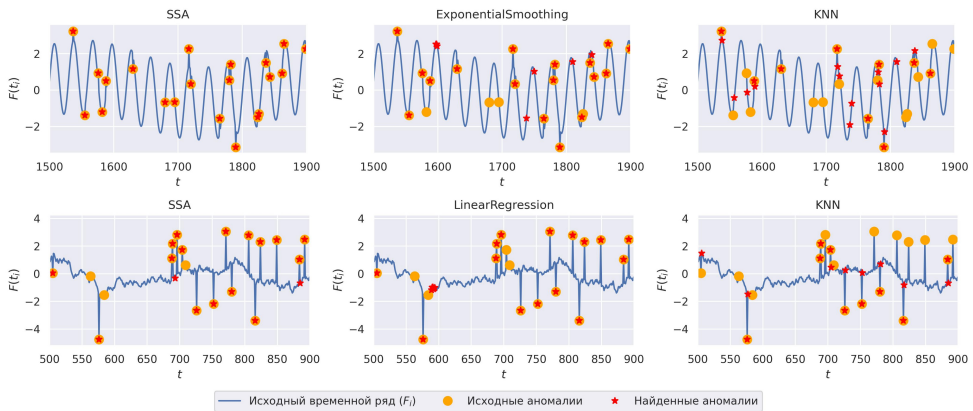


Рис. 1. Результаты работы алгоритмов для суперпозиции гармоник (верхняя строка) и сглаженного шума (нижняя строка)

Таблица 1. Сравнение алгоритмов для суперпозиции гармоник и сглаженного шума

Алгоритм	Суперпозиция гармоник				Сглаженный шум			
	Precision	Recall	F1	Время, с	Precision	Recall	F1	Время, с
SSA	1,000	0,959	0,979	1,571	0,816	0,851	0,833	1,466
Прогнозирование	0,776	0,776	0,776	15,011	0,700	0,745	0,722	0,502
KNN	0,420	0,429	0,424	0,043	0,560	0,583	0,571	0,031

- [1] Banerjee A., Chandola V., Kumar V. Anomaly Detection: A Survey // ACM Computing Surveys. — 2009. — Vol. 41, No. 15. — Pp. 1–58.

- [2] *Губко П., Горчаков А., Буркина М.* Метрики классификации и регрессии // Учебник по машинному обучению: электронный ресурс. — URL: <https://education.yandex.ru/handbook/ml/article/metriki-klassifikacii-i-regressii> (дата обращения: 02.08.2024).
- [3] Nearest Neighbors // User guide: электронный ресурс. — URL: <https://scikit-learn.org/stable/modules/neighbors.html> (дата обращения: 02.08.2024).
- [4] *Голяндина Н. Э.* Метод «Гусеница»-SSA: анализ временных рядов. Учеб. пособие. — СПб.: Изд-во СПбГУ, 2004. — 76 с.

Контролируемая генерация синтетических данных для задачи диаризации

*Майсурадзе Арчил Ивериевич*¹

maysuradze@cs.msu.ru

Попов Дмитрий Николаевич^{1*}

dr.dmitrii2000@yandex.ru

¹Москва, МГУ им. М. В. Ломоносова

Задача диаризации заключается в определении говорящего в каждый момент времени и является актуальной и сложной задачей в области обработки речевых данных. Важным этапом диаризации является сегментация аудиоряда, которая включает выделение сегментов речи одного говорящего и определение интервалов перекрывающейся речи. Современные модели диаризации требуют больших объемов данных с точной разметкой для обучения. Обычно обучение моделей сегментации происходит на реальных наборах данных, полученных при помощи ручной разметки. Это представляет собой дорогостоящий и трудоемкий процесс. В современных диаризационных системах тема синтетических данных либо не затрагивается, либо упоминается без описания процедуры их создания. В данной работе мы предлагаем метод генерации синтетических данных из примеров речи различных говорящих для обучения моделей и демонстрируем его применимость для обучения моделей сегментации аудиоряда.

Предложенный метод позволяет моделировать реалистичные диалоги с контролируемой долей перекрывающейся речи. Метод имитирует различные сценарии смены говорящих, пауз и одновременной речи и включает в себя дополнительные параметры синтеза данных, такие как изменяющиеся акустические условия и фоновые шумы. Для генерации использовались базы данных с примерами голосов отдельных говорящих. Подобная процедура позволяет получить точную разметку временных интервалов тишины, речи, перекрывающейся речи, и смены говорящего.

В ходе экспериментов мы использовали предложенный метод для создания большого объема синтетических данных, на основе которых была обучена модель для сегментации аудиоряда. Для оценки качества модели были использованы популярные наборы данных, содержащие примеры диалогов в различных условиях. Сравнение проводилось с актуальными решениями в данной обла-

сти. По результатам тестирования обученная на предложенных данных модель продемонстрировала высокие показатели диаризационной точности, что свидетельствует об эффективности предложенного метода генерации данных. Данный метод открывает новые возможности для улучшения систем диаризации без необходимости в дорогостоящем и трудоемком процессе получения и разметки аудиоданных требуемого формата.

Аннотации, вводные и заключительные разделы научных статей и их ранжирование по близости смысловому эталону

Михайлов Дмитрий Владимирович^{1*}

mdv74@list.ru

Емельянов Геннадий Мартинович¹

Gennady.Emelyanov@novsu.ru

¹Великий Новгород, Россия, НовГУ

Подборка публикаций по заданной теме подразумевает поиск оптимального порядка работы с первоисточниками от более общего к более специфическому. В работе [1] предложено решение данной задачи на основе семейства нейросетевых языковых моделей BERT. Поскольку анализируемые фрагменты публикаций в [1] — это их аннотации вместе с заголовками, то актуальной здесь будет проблема полноты изложения основного содержания каждой из работ. Один из вариантов ее решения — расширить аннотацию, используя введенную авторами оценку смысловой связности текста. При этом по каждому предложению Ts_j аннотации $\mathbb{T}s_i$ для отвечающего ему многомерного вектора (эмбединга) вычисляется массив значений $\mathbb{C}s_j$ косинусной близости аналогичным векторам остальных предложений аннотации и выбирается «центр масс» — предложение Ts_{\max} с максимальным суммарным значением близости до остальных предложений. Сама смысловая связность аннотации $\mathbb{T}s_i$ формально определяется как

$$\text{cn}(\mathbb{T}s_i) = \max(\mathbb{C}s_{\max}) / (1.0 + \text{std}(\mathbb{C}s_{\max})) , \quad (1)$$

где $\text{std}(\mathbb{C}s_{\max})$ — среднеквадратическое отклонение значения косинусной близости предложения Ts_{\max} остальным предложениям аннотации, $\max(\mathbb{C}s_{\max})$ — максимальное из значений, представленных в массиве $\mathbb{C}s_{\max}$.

Первым шагом вычисляется значение оценки (1) для исходной аннотации, оно принимается за текущее. Далее в аннотацию добавляется то предложение из объединенного множества предложений вводного (*introduction*) и заключительного (*conclusions*) разделов, для которого величина оценки (1) по расширенной аннотации будет максимальной. Если при этом новое значение оценки (1) больше текущего, то на следующей итерации оно становится текущим, а процесс повторяется для объединенного *introduction* и *conclusions*, из которого удаляется только что добавленное в аннотацию предложение. Процесс завершается, когда на очередной итерации новое значение оценки (1) оказывается меньше текущего, а в качестве результата возвращается аннотация из предыдущей итерации.

Реализация предложенного решения на Python 3.10 представлена по адресу: <https://portal.novsu.ru/file/2135027>. При этом рейтинг публикации среди работ коллекции, ассоциируемый с близостью аннотации эталону, определяется произведением оценки (1) и оценки «силы» смысловой связи публикации с остальной коллекцией [1]. Для указанной «силы» используются две независимые оценки, содержательно близкие оценке (1). В случае оценки относительно центров масс в роли $\mathbb{C}s_{\max}$ будет массив значений косинусной близости

вектора центра масс анализируемой аннотации аналогичным векторам центров масс аннотаций остальных работ. При оценке относительно полных текстов массив состоит из значений косинусной близости эмбединга текста анализируемой и соответствующих эмбедингов остальных аннотаций. Для ранжирования относительно заданной языковой модели отбираются расширенные аннотации с рейтингом не ниже, чем у их исходных вариантов. Аннотации остальных работ коллекции при сравнении здесь берутся в исходном (нерасширенном) варианте.

Работа поддержана грантом РФФИ № 19-01-00006.

- [1] *Михайлов Д. В., Емельянов Г. М.* Трансформерные модели BERT, взаимное сходство смыслов коротких текстов и их ранжирование по близости эталону // Тез. докл. 21-й Всерос. конф. ММРО-2023. — М.: Рос. акад. наук, 2023. — С. 159–161.

Использование графов знаний для решения задачи тематической сегментации научных текстов

*Шарафиев Артур Флорович*¹

whiskeydudev@gmail.com

Гаврилов Максим Сергеевич^{1*}

cobraj@yandex.ru

¹Москва, ИПУ РАН

Тематическая сегментация представляет из себя задачу разбиения текста на тематически связанные абзацы, т.е. фрагменты, в которых говорится об одном и том же. Несмотря на то, что эта задача довольно хорошо изучена [1], тематической сегментации с применением графов знаний уделено мало внимания. При этом, графы знаний позволяют учитывать информацию из внешних источников, тем самым, в теории, увеличивая качество сегментации [2].

В данной работе рассматривается ряд подходов к решению этой задачи, используются различные типы внешних графов, а также рассматриваются методы, учитывающие структуру предложений и строящие граф на основе сегментируемого текста. Представленные методы сравниваются друг с другом и с классическими методами, не использующими графы знаний.

Цель данной работы — рассмотреть широкий спектр алгоритмов, применимых для сегментации текстов с помощью графов знаний, составить их классификацию и определить наиболее перспективное направление для дальнейших исследований в данной области.

В работе рассматриваются три типа графов знаний: граф тезауруса, граф классификатора и синтаксический граф.

Рассматриваются два основных подхода к сегментации текстов: сегментация с помощью оптимизации функционала качества, или глобальная сегментация, и сегментация с помощью оценки схожести соседних предложений, или локальная сегментация. Также рассматриваются алгоритмы на основе накопительных графов, которые работают подобно алгоритмам локальной сегментации, однако

вместо сравнения двух предложений, производят сравнение сегмента с соседним предложением.

Была проведена валидация рассмотренных методов на наборе данных, составленном из 65 статей на 27 тем. Результаты представлены в таблице 1. Метрика pk сравнивает сегментации по принадлежности концов скользящего окна к одному и тому же сегменту, или разным сегментам. Метрика windowdiff осуществляет сравнение по количеству разрывов сегментов в скользящем окне. Оба показателя оценивают разницу между представленной и эталонной сегментацией, т. е., чем меньше значение, тем лучше сегментация.

Таблица 1. Показатели качества работы рассмотренных алгоритмов

алгоритм	windowdiff	pk
базовый алгоритм (bert)	0,664	0,343
глобальная сегментация (тезаурус)	0,744	0,617
глобальная сегментация (классификатор)	0,629	0,606
локальная сегментация (тезаурус)	0,623	0,476
локальная сегментация (классификатор)	0,644	0,383
локальная сегментация (синтаксический граф)	0,644	0,383
накопительный граф	0,485	0,404

- [1] *Hearst A. M.* TextTiling, Segmenting Text into Multi-Paragraph Subtopic Passages // Computational Linguistics. — 1997. — Vol. 23. — Pp. 33–64.
- [2] *Das A. Das P. P.* Incorporating Domain Knowledge to Improve Topic Segmentation of Long MOOC Lecture Videos // arXiv. — 2020. — <https://arxiv.org/abs/2012.07589>

Создание набора данных для комбинированной классификации рецензий исходного кода

Петрова Полина Андреевна^{1,2*}

ppolina@ispras.ru

*Марков Сергей Игоревич*¹

markov@ispras.ru

Качанов Владимир Владимирович^{1,3}

vkachanov@ispras.ru

¹Москва, ИСП РАН им. В. П. Иванникова

²Москва, МГУ им. М. В. Ломоносова

³Москва, МФТИ

Рецензирование исходного кода является неотъемлемой частью жизненного цикла разработки программного обеспечения [1]. Под рецензией подразумевается короткое сообщение, содержащее рекомендации по исправлению исходного кода автора. Классификация рецензий помогает проанализировать статистику комментариев и выяснить, насколько эффективно рецензирование. Классификация может сэкономить время на прочтение рецензий и ускорить процесс за-

пуска кода. Также, она помогает решать вспомогательные задачи, например определять полезность комментариев для автора кода. В данной задаче, как и во многих других задачах анализа данных качество классификации в значительной мере зависит от объема и чистоты набора данных. Основной целью работы было создать такой набор данных и протестировать на нем различные методы классификации рецензий.

Для тестирования и оценки методов требуется достаточно большой датасет. Мы объединили четыре датасета, над каждым из которых работали группы исследователей [2–5], изучавшие данную тему. А также разметили вместе с командой вручную 3200 комментариев. Итоговый датасет из 9645 комментариев выложен в открытый доступ, для того чтобы другие команды могли продолжить изучать данную задачу. Нами создана иерархическая классификация на основе уже имеющихся датасетов: итоговый датасет разделен на 16 классов, и из них составлено 5 групп.

В рамках исследования были реализованы различные методы для проведения классификации, проведено сравнение классификаторов по соотношению качество/количество потребляемых ресурсов. Для этого мы использовали векторные представления: `CountVectorizer`, `TfidfVectorizer`; различные классификаторы, среди них: классические — `Decision Tree`, `LogReg`, `SVM`; ансамблевые — `Random Forest`, `Gradient-Boosted Decision Trees`, `eXtreme Gradient Boosting`, `Categorical Boosting`. Также были проведены эксперименты на более сложных моделях: `Word2Vec`, `FastText`, `GloVe` и разновидности `BERT`.

Комбинируя между собой эмбединги и классификаторы, мы получили наиболее эффективный метод для классификации рецензий исходного кода. В соотношении качество/количество потребляемых ресурсов неплохой результат показал `FastText`, с его скоростью работы можно запускать на CPU, $f1 - macro = 0,579$. Лучшим классификатором оказался `CodeBERT` [6], $f1 - macro = 0,775$. Наше исследование показывает возможность классификации рецензий исходного кода с приемлемой точностью. Это поможет выявлять обсуждения, требующие дополнительного внимания разработчиков проекта.

- [1] *Bacchelli A., Bird C.* Expectations, Outcomes, and Challenges of Modern Code Review // 2013 35th Int. Conf. Software Engineering (ICSE). — San Francisco, CA, USA, 2013. — Pp. 712–721. <https://doi.org/10.1109/ICSE.2013.6606617>
- [2] *Li Z., Yu Y., Yin G., Wang T., Fan Q., Wang H.* Automatic Classification of Review Comments in Pull-Based Development Model // 29th Int. Conf. Software Engineering and Knowledge Engineering. — 2017. <https://doi.org/10.18293/SEKE2017-039>
- [3] *Muntazir F.*, Dissecting GitHub Code Reviews: A Text Classification Experiment // [Электронный ресурс] — <https://mfadhel.com/github-code-reviews/> (дата обращения: 03.04.2024)
- [4] *Ochodek M., Staron M., Meding W., Söder O.* Automated Code Review Comment Classification to Improve Modern Code Reviews // Springer Nature Switzerland AG

- 2022 D. Mendez et al. (Eds.) SWQD 2022. — Springer, Cham, 2022. — Vol. 439. — Pp. 23–40.
- [5] *Turzo A., Faysal F., Poddar O., Sarker J., Iqbal A., Bosu A.* Towards Automated Classification of Code Review Feedback to Support Analytics // arXiv. — 2023. — <https://arxiv.org/abs/2307.03852>
- [6] *Feng Z., Guo D., Tang D., Duan N., Feng X., Gong M., Shou L., Qin B., Liu T., Jiang D., Zhou M.* CodeBERT: A Pre-Trained Model for Programming and Natural Languages // arXiv. — 2020. — <https://arxiv.org/abs/2002.08155>

Мастерская знаний: большие языковые модели для поиска и систематизации научной информации

*Воронцов Константин Вячеславович*¹*

k.vorontsov@iai.msu.ru

¹Москва, ФИЦ «Информатика и управление» РАН

Большие языковые модели (Large Language Model, LLM) на основе контекстно-зависимых векторных представлений слов позволяют сегодня решать задачи анализа текстов с качеством, недостижимым еще несколько лет назад.

Развивая опыт построения русскоязычной модели научных текстов [1], мы предлагаем двуязычную модель SciRus-tiny (MSU), стремясь минимизировать ее размер и вычислительные затраты при обучении и применении. Модель имеет архитектуру RoBERTa [2] и обучается сначала восстановлению маскированных токенов по 44М аннотаций научных статей на русском и английском языках (15В токенов), затем по контрастным данным о цитировании и двуязычным аннотациям. Эта модель опережает лучшую модель англоязычных научных текстов SciNCL на 13 из 24 задачах открытого бенчмарка SciRepEval, имея при этом почти в 5 раз меньше параметров (23М против 110М), вдвое меньшую размерность векторных представлений (312 против 768) и вдвое больший контекст (1024 токенов против 512). Для тестирования функций мультязычного поиска и классификации мы построили бенчмарк ruSciBench, состоящий из 7 задач (14 метрик) на 400К параллельных аннотациях на русском и английском языках. На этом бенчмарке модель SciRus-tiny (MSU) лишь немного уступает моделям, превосходящим ее по числу параметров в 20 раз и более.

Модель SciRus-tiny (MSU) внедрена в систему поиска семантически схожих публикаций научной электронной библиотеки eLibrary.ru. Также на ее основе мы разрабатываем поисково-рекомендательную систему «Мастерская Знаний», прототип которой доступен по адресу <https://scisearch.ru>. Система позволяет пользователю собирать тематические подборки научных статей с помощью семантического векторного поиска и в дальнейшем анализировать свои подборки (возможно, расширенные результатами поиска) с помощью линейки сервисов.

Сервис мониторинга — для рекомендации добавления в пользовательские подборки новых публикаций по мере их появления в открытом доступе.

Сервис систематизации знаний с помощью вероятностной тематической модели с автоматическим именованием и суммаризацией тем [3, 4].

Сервис полуавтоматического реферирования — для ускорения написания научных обзоров по подборкам. Сервис сначала рекомендует пользователю, в каком порядке упоминать статьи подборки, затем для каждой статьи ранжирует фразы, релевантные выбираемым пользователем аспектам статьи.

Сервис образовательной траектории — для рекомендации порядка чтения документов подборки с учетом широты тематики и когнитивной сложности [5].

Сервис анализа трендов — для определения моментов появления растущих научных трендов как новых тем с характерной устойчивой терминологией [6].

Сервис картирования знаний — для выделения в подборке основных идей, наиболее важных для понимания, усвоения и применения знаний, путем разделения каждой идеи на основные подыдеи, до заданного уровня детализации, с помощью промптинга генеративных моделей, иерархической суммаризации и текстографической визуализации в виде интеллект-карт (mind map) [7].

- [1] Герасименко Н. А., Чернявский А. С., Никифорова М. А. RuSciBERT: языковая модель на базе архитектуры трансформера для получения семантических векторных представлений научных текстов на русском языке // Доклады РАН. Математика, информатика, процессы управления. — 2022. — Т. 508, № 1. — С. 104–105.
- [2] Yinhan Liu et al. RoBERTa: A Robustly Optimized BERT Pretraining Approach // arXiv. — 2019. — <https://arxiv.org/abs/1907.11692>
- [3] Vorontsov K. V. Rethinking Probabilistic Topic Modeling from the Point of View of Classical non-Bayesian Regularization // Data Analysis and Optimization. — Springer, 2023. — Pp. 397–422.
- [4] Герасименко Н. А., Потапова П. С., Янина А. О., Воронцов К. В. Применение вероятностного тематического моделирования в четырех задачах разведочного информационного поиска // Информационный бюллетень Российской библиотечной ассоциации. — 2022. — № 98. — С. 43–48.
- [5] Veselov A. S., Ereemeev M. A., Vorontsov K. V. Estimating Cognitive Text Complexity with Aggregation of Quantile-based Models // Computational Linguistics and Intellectual Technologies. — Dialogue, 2023. — Pp. 525–538.
- [6] Герасименко Н. А., Чернявский А. С., Никифорова М. А., Никитин М. Д., Воронцов К. В. Инкрементальное обучение тематических моделей для поиска трендовых тем в научных публикациях // Доклады РАН. Математика, информатика, процессы управления. — 2022. — Т. 508, № 1. — С. 106–108.
- [7] Воронцов К. В., Курилов В. А. Карты знаний. На пути к доверенным языковым моделям и системам представления знаний // BIS Journal — Информационная безопасность банков. — 2024. — № 3 (54), С. 28–33.

Применение моделей детекции объектов для анализа видеопотоков в реальном времени

*Мартыненко Александр Станиславович*¹

s.martynenko@bk.ru

*Миromanов Даниил Борисович*¹

kononov.miromanov@gmail.com

*Семенов Михаил Евгеньевич*¹

semvon.me@talantiuspech.ru

Хубиев Касымхан Юсуфович^{1*}

kasymkhankhubievnis@gmail.com

¹п. г. т. Сириус, Научно-технологический университет «Сириус»

Управление очередью становится все более важной задачей, особенно с учетом растущего использования роботов-доставщиков, и камеры видеонаблюдения в общественных местах могут помочь в анализе и управлении очередью через потоковую видеоаналитику.

Для проведения экспериментов мы использовали видеопоток с двух камер наблюдения (фронтальная, тыловая) в столовой, который передавался на сервер по RTSP протоколу, где далее подвергался раскадровке. На каждом кадре были выделены следующие полигоны: области очереди, посадочных мест, задачи, рабочей зоны. Объекты разделялись на следующие классы: люди в очереди, сотрудники и посетители столовой. В работе были рассмотрены: модель рекуррентной нейронной сети Fast R-CNN [1], трансформерная модель DETR [2] и YOLO v8 [3]. Для обучения моделей мы использовали видеокарту Nvidia A100 с объемом видеопамяти 40 ГБ.

Для оценки качества моделей были применены следующие метрики: Precision, Recall, mAP0.5, mAP0.5:0.95, Latency. Эксперименты показали (см. табл. 1), что трансформерная модель DETR достигает наилучшего качества детекции людей в очереди в случаях наложения двух объектов, но уступает по скорости двухстадийной модели YOLO v8; модель Fast R-CNN показала наихудшие результаты детекции. В результате была выбрана модель YOLO v8, как оптимальная модель по качеству и скорости.

Таблица 1. Результаты экспериментов

Модель	Precision	Recall	mAP0.5	mAP0.5:0.95	Latency (с)
Fast R-CNN	0,62	0,71	0,41	0,34	44
YOLO v8	0,82	0,65	0,76	0,53	34
DETR	0,84	0,71	0,78	0,55	40

Дообучение (fine tuning) модели на трех классах (сотрудник, клиент в очереди, клиент в зале) уменьшило время обработки кадров в три раза. Разработанный программно-аппаратный комплекс определяет длину очереди и загруженность зала в реальном времени. Дальнейшее развитие проекта направлено на оптимизацию процесса встраивания программы управления очередью на производственных объектах в формате plug-and-play.

- [1] *Girshick R.* Fast R-CNN // 2015 IEEE Int. Conf. Computer Vision (ICCV), Santiago, Chile. — 2015. — Pp. 1440–1448.
- [2] *Carion N., Massa F., Synnaeve G., Usunier N., Kirillov A., Zagoruyko S.* End-to-End Object Detection with Transformers // Vedaldi A., Bischof H., Brox T., Frahm JM. (eds) Computer Vision — ECCV 2020. Lecture Notes in Computer Science. — Springer, Cham., 2020. — Vol. 12346. — Pp. 213–229.
- [3] *Redmon J., Divvala S., Girshick R., Farhadi A.* You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. — 2016. — P. 779–788.

Задачи оптимизации использования распределенной облачной вычислительной инфраструктуры

Писковский Виктор Олегович^{1,2,*}

vpiskovski@lvk.cs.msu.ru

*Грушо Александр Александрович*²

grusho@yandex.ru

*Забезжайло Михаил Иванович*²

m.zabezhailo@yandex.ru

¹Москва, ВМК МГУ им. М.В. Ломоносова

²Москва, ФИЦ ИУ РАН

В докладе представлены результаты решения задач оптимального использования распределенной облачной вычислительной инфраструктуры. В качестве таких задач рассматривались:

- прогнозирование качества получаемой клиентом сетевой услуги по имеющимся трассам работы оборудования при реализации сервисов,
- оценка времени выполнения вычислительных задач центром обработки данных по журналу регистрации работ за 2023 год системы управления заданиями высокопроизводительного вычислителя SLURM,
- время активации образа виртуальной машины в качестве абонентского облачного терминала, загружаемого с общего разделяемого сетевого ресурса.

В рамках первой задачи [1–4] решалась проблема нахождения оптимального метода по имеющимся в распоряжении наборе данных:

$$F_k(\vec{x}) = \arg \min_{F_k} \|F_k(\vec{x}) - y_k\|_{l_2}, \quad \vec{x} \in \mathbb{R}^D, \quad \{\vec{x}\}_{n=1}^N, \quad N \sim 20\,000,$$

где F_k — искомый метод, \mathbb{R}^D — пространство показателей, y_k — прогнозируемые характеристики, N — количество трасс, используемых для обучения модели.

В качестве набора данных были использованы трассы Центра обработки данных Шведского Королевского технологического университета (Стокгольм, Швеция) [5] для сервисов видео по запросу и СУБД общим объемом десятки тысяч записей (N) по полторы тысячи показателей (D) в каждой.

В результате исследований были отобраны методы:

- на основе случайного леса (Random Forest, RF),

- анализ главных компонент (Principal Component Analysis, PCA),
- усеченного сингулярного разложения (Truncated Singular Value Decomposition, TSVD),
- на основе нейронных сетей: многослойный перцептрон (Multilevel Perceptron, MLP) и сверточная сеть (Convolution Neuron Network, CNN).

Наилучшие результаты получены для композиции метода анализа главных компонент с методами на основе случайного леса или нейронных сетей.

Результат решения первой задачи был применен ко второй проблеме оценки времени выполнения вычислительных задач Центром обработки данных, показав вполне приемлемую точность прогноза.

В рамках третьей задачи [6] исследовались время старта виртуальной машины (ВМ) в режиме расширенного фильтра записи в зависимости от применяемых сетевых характеристик, протоколов доступа и файловых систем. Работа ВМ осуществляется непосредственно из сетевого хранилища, доступного только в режиме чтения. В результате разработаны и экспериментально обоснованы практические рекомендации по оптимальной организации работы с ВМ в рассматриваемых условиях.

В планах работ — расширить комплекс методов техниками матричного разложения для восстановления пропущенных данных и понижения размерности задачи, применить промышленные отечественные платформы машинного обучения ГосНИИАС «Plat2023»/«Платформа-ГНС» [7], использовать методы сжатия информации для уменьшения трафика при обмене данными между узлами инфраструктуры и внедрения методов идентификации анализируемых данных.

- [1] *Лычева Е. О., Писковский В. О.* Прогноз выполнения задач VOD и KV БД на оборудовании сетей доступа // Ломоносовские чтения. Научная конференция. 20 марта — 3 апреля 2024 г.: тезисы докладов. — М.: МАКС Пресс, 2024. — С. 119–120.
- [2] *Могиленец В. М., Писковский В. О.* Применение CNN для прогноза выполнения задач VoD и KV БД на оборудовании сетей доступа // Ломоносовские чтения. Научная конференция. 20 марта — 3 апреля 2024 г.: тезисы докладов. — М.: МАКС Пресс, 2024. — С. 121–122.
- [3] *Писковский В. О.* Прогнозирование временных характеристик периферийных вычислений при реализации прикладных сетевых сервисов // Международная конференция «Математика в созвездии наук». — М.: Изд-во Московского университета, 2024. — 815 с. ил., стр. 639–640.
- [4] *Писковский В. О., Грушо А. А.* Прогнозирование временных характеристик прикладных сетевых сервисов // Труды XIV Всероссийского совещания по проблемам управления (ВСПУ-2024, Совещание), 17–20 июня 2024 года, ИПУ РАН (г. Москва, Россия). (Срок выхода публикации — октябрь 2024 г.)
- [5] Data Traces from a Data Center Testbed // [Электронный ресурс] — www.kaggle.com/datasets/jaliltaghia/data-traces-from-a-data-center-testbed (дата обращения: 11.09.2023)

- [6] *Сагалевич В. Д., Писковский В. О.* Исследование времени подготовки АРМ к работе при использовании разделяемого сетевого репозитория образов ВМ // Ломоносовские чтения. Научная конференция. 20 марта — 3 апреля 2024 г.: тезисы докладов. — М.: МАКС Пресс, 2024. — С. 122–123.
- [7] Унифицированная программная платформа для разработки конечно ориентированных программных комплексов автоматического распознавания объектов на основе нейросетевых подходов «Платформа-ГНС» // [Электронный ресурс] — <https://www.gosniias.ru/platform.html> (дата обращения: 03.04.2024)

О концепции построения универсальной системы городской среды

Рудикова-Фронцёфер Лада Владимировна^{1*}

lada.rudikowa@gmail.com

*Семенчук Наталья Владимировна*¹

senata155@gmail.com

*Дейцева Анна Геннадьевна*¹

dejceva_ag@grsu.by

¹Гродно, Гродненский государственный университет имени Янки Купалы

В работе формулируются требования к построению общей концепции системы городской среды; выявлены ключевые аспекты, которые влияют на разработку общей архитектуры универсальной системы; разработаны общая модель данных и функциональная модель для системы; определен общий подход к разработке программного обеспечения на основе технологии складирования данных.

Первоначально в системе необходимо реализовать возможность постоянного поступления в базу информации обо всех городах Беларуси [1]. Зная информацию о городе, можно получить агрегированные данные о количестве объектов городской среды разного типа, возможных тенденциях социальной городской среды, трендах в активности населения и т. п. По полученным и общим статистическим данным можно рассчитывать индекс городской среды и населенных пунктов. Также одной из важнейших задач разработки универсальной системы является предоставление удобного и понятного пользовательского интерфейса, предоставление пользователям необходимых данных в соответствии с заданными критериями отбора, обработанных данных и необходимых визуализаций. Такой подход позволит иметь представление о текущей ситуации в городах и населенных пунктах, проводить необходимый анализ данных и выстраивать стратегию для развития и улучшения благосостояния городской среды.

Можно перечислить основные функциональные требования, которые должны поддерживаться универсальной системой сбора и анализа данных городской среды: получение основной информации о годах и общей статистике; получение данных в соответствии с заданными критериями поиска о городах Беларуси; добавление и модификация данных, связанных с городской средой; определение индекса города для крупных городов Беларуси; просмотр карты Беларуси

с отображением статуса городов в соответствии с полученным значением индекса города; просмотр подробной информации о городе и городских объектах; визуализация требуемых данных на карте; анализ требуемых данных и построение трендов; возможность получения агрегированных данных по различным критериям и т. п.

Отметим, что перечисленные выше основные функции позволят пользователям узнать необходимую информацию о состоянии города и качестве жизни, сравнить несколько городов по определенным параметрам. Предполагается, что концепция универсальной системы накопления и анализа данных городской среды будет базироваться на основе технологии складирования данных. В основе которой лежит многоуровневая архитектура реализации с распределенными узлами доступа к данным. Кроме того, в системе городской среды предусмотрена подсистема интеллектуального анализа данных, которая будет поддерживать принятие решений в аспекте развития городских территорий и процессов. Указанная подсистема интеллектуального анализа данных использует следующие методы Data Mining: деревья решений, алгоритмы кластеризации, в том числе и масштабируемые и алгоритмы обнаружения ассоциативных связей между событиями.

- [1] *Mityagin S., Yakimuk I., Rudikowa L., Myslivec O., Drozhzhin A.* City Information Modeling: The system Approach for Formation Requirement in Spatial Development // *Procedia Computer Science.* — 2020. — Vol. 178. — Pp. 134–144.

Визуальная навигация беспилотного летательного аппарата на основе анализа изображений подстилающей поверхности

*Ермаков Павел Григорьевич*¹★

ErmakovPG@gosniias.ru

*Лебедев Максим Алексеевич*¹

MLebedev@gosniias.ru

*Фурс Егор Васильевич*¹

Furs_Egor@gosniias.ru

¹Москва, ФАУ «ГосНИИАС»

Сегодня малые беспилотные летательные аппараты (БЛА) со взлетной массой менее 100 кг широко используются как для решения гражданских задач: поиск и спасение людей, тушение пожаров или доставка грузов в отдаленные районы, так и в военных целях: выполнение целеуказания, осуществление разведки и т. д. В условиях отсутствия сигнала спутниковых навигационных систем БЛА должен определять свое местоположение при помощи имеющихся бортовых средств [1].

В данной работе решается задача визуальной навигации БЛА на основе анализа подстилающей поверхности при помощи заложенного на этапе подготовки полетного задания набора спутниковых снимков (в том числе полученных в другое время года и имеющих отличия от актуальной сцены) в условиях ограниченных бортовых вычислительных ресурсов. Задача решается при помощи

методов на основе детектирования ключевых точек с последующим их сопоставлением и анализом. В работе описываются этапы реализации алгоритма, осуществляется сравнение использования различных методов детектирования ключевых точек (SIFT, ORB, BRISK, AKAZE), приводится анализ быстродействия и результаты тестирования с реальных полетов БЛА.

По результатам имитационного моделирования процесса функционирования решения задачи визуальной навигации на основе использования набора спутниковых изображений подстилающей поверхности (общедоступный сервис Яндекс Карты) можно утверждать следующее:

1. применение SIFT на предварительно сегментированном зимнем изображении подстилающей поверхности позволяет повысить количество детектируемых ключевых точек по сравнению с другими рассмотренными методами в данном исследовании;

2. более робастным методом детектирования ключевых точек к углу поворота изображения подстилающей поверхности является SIFT;

3. вычислительное время определения географических координат носителя составило 37 мс для 150 ключевых точек (Intel® Core™ i7-8565U).

- [1] Али Б., Садеков Р. Н., Цодокова В. В. Алгоритмы навигации БПЛА с использованием систем технического зрения // Гирскопия и навигация. — 2022. — Т. 30, № 4 (119). — С. 87–105. <https://doi.org/10.17285/0869-7035.00105>

О построении системы считывания информации с экрана контроллера

Дожукин Александр Александрович¹*

dalex@ccas.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» Российской академии наук

При разработке новых контроллеров возникает задача их тестирования. Контроллер — это устройство с ограниченными ресурсами, способное управлять исполнительными механизмами и реализующее сложную логику. Немалая часть тестирования их работы ложится на людей, поэтому вызывает интерес автоматизация рутинных операций, важной частью которой является считывание информации непосредственно с экрана устройства.

Прототип системы разрабатывается для контроллеров Smartweb-K [1] с монохромным экраном 128 × 64 точек. Изображение на экране определяется выбранным меню и состоянием контроллера. Тип меню задает расположение текстовых и графических элементов (см. Рис. 1), а состояние — их содержимое. Для прототипа выбраны 29 основных типов меню, что связано с трудоемкостью получения обучающей выборки. Описания каждого предоставлены разработчиками контроллера. Задача естественным образом разбивается на две: определить текущий тип меню и считать текст и иконки в соответствующих ему позициях.

Для первого этапа подобрана трехслойная сверточная сеть [2, 3]: $(32, 3 \times 3)$, $(16, 3 \times 3)$, $(29, 60 \times 124)$, где числа показывают число выходных каналов и размер ядра. В результате применения сети к исходному изображению получается 29-мерный вектор. Сеть обучается так, чтобы вектор содержал оценки за типы меню. Для создания обучающей выборки со случайным заполнением полей была написана программа (см. Рис. 1). Сеть очень быстро обучается и работа-



Рис. 1. Примеры искусственных изображений экрана

ет, и позволяет одновременно определять тип меню и находить его расположение на фоне, что полезно при неточном положении камеры. Сеть, примененная к изображению большего размера, вместо вектора выдает тензор, в котором каждому возможному расположению экрана соответствует тот же 29-мерный вектор. Положение глобального максимума соответствует типу меню и координате экрана, что подтвердилось в 100 % тестов.

На втором этапе хотелось сохранить преимущества использования библиотеки `pytorch` [2]. Эксперименты проводились с одним из используемых в контроллере шрифтов (143 символа), также предоставленном разработчиками. Для работы с отдельными буквами предыдущая архитектура не подошла. Вместо этого предложено использовать аппарат сверточного слоя для непосредственного сравнения изображения с масками букв. Слой реализует свертку $(2X - 1) \otimes T$, где X — изображение; T — маска буквы (1 — буква, -1 — фон и отступы, 0 — за пределами отступов), нормированная на ширину. Аналогично первому этапу это позволяет одновременно определять положение и класс буквы.

На тесте с одиночным символом, случайно помещенном в строке, схема дала точность 99,3 %. На тесте со случайным сочетанием 5 букв подряд точность составила около 97 %. Например, в последовательности «ВмыЧ,», уверенно нашлась буква «ь», что дало «ВмыЧ,» в качестве результата. Для точного решения этой задачи требуется привлечение контекста.

- [1] SmartWEB-K // [Электронный ресурс] — <https://www.teplostart.ru/download/booklets/SmartWEB-K.pdf> (дата обращения: 04.08.2024)
- [2] Le Cun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D. Handwritten Digit Recognition with a Back-Propagation Network // NIPS'89: Proc. 2nd Int. Conf. Neural Information Processing Systems. — 1989. — Pp. 396–404.
- [3] torch.nn.Conv2d // [Электронный ресурс] — <https://pytorch.org/docs/stable/generated/torch.nn.Conv2d.html> (дата обращения: 04.08.2024)

Обеспечение ситуационной осведомленности экипажа сверхзвуковых гражданских самолетов

*Фурс Егор Васильевич*¹*

Furs_Egor@gosniias.ru

*Воробьев Тимур Константинович*¹

VorobievTK@gosniias.ru

*Зотов Андрей Владимирович*¹

ZotovAV@gosniias.ru

*Комаров Денис Валерьевич*¹

DKomarov@gosniias.ru

¹Москва, ФАУ «ГосНИИАС»

Разработка новых сверхзвуковых гражданских самолетов (СГС) требует поиска и отработки новых технических решений по аэродинамической компоновке, силовой установке и конструктивно-силовой схеме, обеспечивающих высокий уровень технико-экономических показателей самолета [1]. Все это приводит к значительному ограничению кругового обзора для экипажа и уменьшению его ситуационной осведомленности. В рамках данной работы предлагается решать задачу обеспечения ситуационной осведомленности экипажа СГС при помощи внедрения системы внешнего видения (СВВ), которая представляет собой комплексное программно-аппаратное решение, обеспечивающее интеллектуальную информационную поддержку экипажа СГС. СВВ состоит из многоспектральной оптико-электронной системы (МОЭС), бортовых высокопроизводительных вычислителей и функционального программного обеспечения (ФПО).

МОЭС представляет из себя два функциональных контура: контур курсового канала и контур обзорного канала. Курсовой канал обеспечивает выдачу изображений подстилающей поверхности при различных метеорологических условиях за счет применения 3 камер различного спектрального диапазона. Обзорный канал обеспечивает формирование панорамного изображения видимого диапазона для всестороннего обзора закабинного пространства при осуществлении руления по аэродромному полю.

К бортовым высокопроизводительным вычислителям предъявляются повышенные требования по виброустойчивости, помехозащищенности, энергопотреблению и теплоотводу. Современное ФПО базируется на аппарате глубоких нейронных сетей (ГНС), позволяющих в автоматическом режиме решать следующие задачи (в том числе в сложных метеорологических условиях): улучшение визуального качества изображений закабинной обстановки, детектирование взлетно-посадочной полосы, детектирование линий разметки рулежных дорожек, распознавание и классификация уровня угрозы препятствий при движении воздушного судна по аэродрому. Стоит также отметить, что использование аппарата ГНС накладывает высокие требования: к производительности вычислителей, к использованию в них многоядерных процессоров, к наличию графических процессоров с большим объемом видеопамати.

- [1] *Сидняев Н. И.* Перспективы развития современных сверхзвуковых пассажирских самолетов // Инженерный журнал: наука и инновации. — 2023. — №.1. <https://doi.org/10.18698/2308-6033-2023-1-2243>

Программный комплекс для анализа и прогнозирования временных рядов метеоданных

Арзангулян Марк Эмилевич¹*

arzangulian.me19@physics.msu.ru

¹Москва, МГУ имени М. В. Ломоносова, Физический факультет

Современные методы анализа временных рядов играют ключевую роль в прогнозировании метеорологических данных. В данной работе представлен уникальный программный комплекс с открытым исходным кодом, объединяющий математические и нейросетевые подходы для обработки временных рядов метеоданных. Комплекс отличается от аналогичных решений высокой гибкостью и возможностью интеграции различных методов, включая как традиционные статистические, так и передовые нейросетевые подходы, что делает его универсальным инструментом для широкого круга пользователей — от новичков до опытных исследователей. Основное внимание уделено интеграции традиционных методов, таких как ARMA, ARIMA, SARIMA, вейвлет-преобразование, преобразование Фурье и морфологический анализ, с нейросетевыми моделями, такими как Long Short-Term Memory.

Комплекс разработан на языке Python с использованием публичных пакетов и имеет открытый исходный код, обеспечивая доступность и возможность модификации. Программное обеспечение может быть запущено как локально, так и в облаке, что позволяет пользователям работать с ним через интернет. Более того, комплекс доступен онлайн по ссылке: <https://time-series-msu-ff.streamlit.app>, что обеспечивает мгновенный доступ к его функциональности без необходимости установки.

Важной особенностью комплекса является возможность комбинированного анализа, в котором результаты одного метода могут быть использованы в качестве входных данных для другого, повышая эффективность и точность прогнозирования.

Комплекс включает модули для предобработки данных, статистического анализа, выделения трендов и сезонных компонентов, а также прогнозирования различными методами. Интуитивно понятный интерфейс делает его доступным как для специалистов, так и для начинающих.

Разработанный с учетом потребностей исследователей и практиков, комплекс подходит для анализа и прогнозирования временных рядов в различных областях. Возможность расширения и адаптации делает его универсальным решением для пользователей.

Таким образом, представленный программный комплекс является эффективным и гибким инструментом для обработки, анализа и прогнозирования временных рядов, предоставляющим уникальное сочетание открытости, широкого спектра методов и возможностей кастомизации, а также удобство онлайн-доступа.

- [1] *Дещеревский А. В., Журавлев В. И., Никольский А. Н.* WINABD — Пакет программ для сопровождения и анализа данных геофизического мониторинга // Институт физики земли РАН, 2023.
- [2] *Faniband Y. P., Ishak I., Sait S. M.* A Review of Open Source Software Tools for Time Series Analysis // arXiv. — 2022. — <https://arxiv.org/abs/2203.0519>

Сегментация и нормализация семантической разметки крыш зданий по аэрофотосъемке для реконструкции трехмерных моделей районов жилой застройки

Мошканцев Петр Владиславович^{1,★}

`petr_mosh@gosniias.ru`

*Князь Владимир Владимирович*¹

`vl.kniaz@gosniias.ru`

Новиков Даниил Иванович^{1,2}

`daninov@gosniias.ru`

Белов Александр Степанович^{1,3}

`the_anskiy@gosniias.ru`

¹Москва, ФАУ ГосНИИАС

²Москва, МГТУ им. Н. Э. Баумана

³Москва, МАИ

Перед нами стояла задача реализации алгоритма сегментации и нормализации разметки зданий для воссоздания трехмерной модели района низково-сотной загородной застройки на основе изображений карт местности, которые не были ранее размечены в каком-либо геоинформационном ресурсе. В первом приближении требовалось ознакомиться с текущими реализациями автоматического восстановления городов по картам, продумать и проработать собственный алгоритм сегментации изображения для получения векторной разметки и нормализации полученной разметки.

В докладе представлен алгоритм создания трехмерной модели района жилой застройки. Отличительной чертой разработанного алгоритма является упрощение растровой разметки, формируемой нейронной сетью Mask R-CNN, для получения векторной прямолинейной разметки, характерной для контуров зданий. Упрощение растровой разметки производится с помощью улучшенного алгоритма нормализации. Векторная прямолинейная разметка контуров зданий в дальнейшем используется для воссоздания трехмерной модели района низково-сотной загородной застройки.

Для работы метода необходима цветная карта района местности разрешением 300×300 пикселей. В качестве отправной точки для сегментации в виде векторной разметки была выбрана конволюционная нейронная сеть Mask R-CNN [5]. Она обрабатывает участки карты и выдает для этих участков семантическую разметку карты местности на единственный класс — здание. В качестве обучающего набора использовалась выборка из 280 741 фрагментов (в виде RGB изображений размером 300×300 пикселей) спутниковых снимков с соответствующими аннотациями [6]. Для оптимизации трехмерной формы строений

используется алгоритм векторизации с последующей нормализацией разметки крыш. В качестве основы для алгоритма нормализации использован алгоритм Рамера и Дугласа-Пекера [2, 4]. Для повышения качества нормализации разметки и для лучшего контурного совпадения разметки и зданий добавлена векторная составляющая направления размещения объекта. Вектор направления используется для подкрепления при нормализации векторной разметки вдоль опорных векторов.

Суть разработанного метода заключается в формировании векторной разметки экземпляров зданий на космофотоснимке. Тестирование алгоритма производилось на независимой тестовой выборке, объемом 200 кадров. В качестве метрики оценивания использовался коэффициент Жаккара между областью предсказанной и истинной разметки. Также для оценки сходства контуров моделей с реальной разметкой использовалась метрика начального расстояния Фреше (Freche Inception Distance). Результаты тестирования показали, что предложенный метод показывает точность сегментации аналогичную точности оригинальной Mask R-CNN по коэффициенту Жаккара. Оценка реалистичности полученных контуров показывает прирост правдоподобности полученных контуров зданий по расстоянию Фреше.

- [1] *Gruen A., Wang X.* CyberCity Modeler, a Tool for Interactive 3-D City Model Generation // Photogrammetric Week '99'. — Heidelberg: Wichmann Verlag, 1999.
- [2] *Douglas D. H., Peucker T. K.* Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature // The Canadian Cartographer. — 1973. — Vol. 10, No. 2. — Pp. 112–122.
- [3] *Zakhor A., Früh C.* Fast 3D Model Generation in Urban Environments // Int. Conf. Multisensor Fusion and Integration for Intelligent Systems. — Baden-Baden, 2001. — Pp. 165–170.
- [4] *Ramer U.* An Iterative Procedure for the Polygonal Approximation of Plane Curves // Computer Graphics and Image Processing. — New York, 1973. — Vol. 1, No. 3. — Pp. 244–256.
- [5] *He K., Gkioxari G., Dollár P., Girshick R.* Mask R-CNN // arXiv. — 2018. — <https://arxiv.org/abs/1703.06870>
- [6] *Mohanty Sh. P., Czakon J., Kaczmarek K. A., Pyskir A., Tarasiewicz P., Kunwar S., Rohrbach J., Luo D., Prasad M., Fleer S. et al.* Deep Learning for Understanding Satellite Imagery: An Experimental Survey // Frontiers in Artificial Intelligence. — Frontiers Media SA, 2020. — Vol. 3.

Обобщение AUC-диаграмм для автоматического поиска признаков болезни Паркинсона и эссенциального тремора

Сушкова Ольга Сергеевна^{1*}

o.sushkova@mail.ru

*Морозов Алексей Александрович*¹

morozov@cplire.ru

*Хохлова Маргарита Николаевна*¹

margokhokhlova@gmail.com

*Кершнер Иван Андреевич*¹

ivan.kershner@gmail.com

*Габова Александра Васильевна*²

agabova@yandex.ru

*Чигалейчик Лариса Анатольевна*³

chigalei4ick.lar@yandex.ru

*Карбанов Алексей Вячеславович*³

doctor.karabanov@mail.ru

¹Москва, Институт радиотехники и электроники им. В. А. Котельникова РАН

²Москва, Институт высшей нервной деятельности и нейрофизиологии РАН

³Москва, ФГБНУ «Научный центр неврологии»

Исследованы и разработаны методы автоматического поиска признаков нейродегенеративных заболеваний «болезнь Паркинсона» и «эссенциальный тремор» на основе AUC-диаграмм и алгоритмов оптимизации. AUC-диаграммы являются новым методом статистического анализа биомедицинских сигналов, основанным на визуализации параметров всплескообразной электрической активности мозга и мышц. Эффективность этого метода была продемонстрирована при решении задач ранней и дифференциальной диагностики болезни Паркинсона и эссенциального тремора. Недостатком данного метода является необходимость построения и анализа большого количества графических диаграмм. В связи с этим, автоматизация анализа AUC-диаграмм является актуальной задачей.

Математическая задача поиска признаков на основе анализа AUC-диаграмм сводится к задаче оптимизации в многомерном пространстве признаков [1]. Отличительной особенностью пространства признаков, построенного с использованием AUC-диаграмм, является наличие сравнительно больших компактных областей, содержащих локальные максимумы и минимумы. Это свойство пространства признаков облегчает поиск решений задачи оптимизации, но при этом требует выбор алгоритмов оптимизации и целевых функций, повышающих вероятность обнаружения глобальных экстремумов. В данной работе исследованы и разработаны методы автоматического поиска глобальных экстремумов в многомерном пространстве признаков всплескообразной электрической активности.

Исследование выполнено за счет гранта Российского научного фонда № 22-75-10079, <https://rscf.ru/project/22-75-10079/>.

- [1] *Sushkova O. S.; Morozov A. A.; Khokhlova M. N.; Kershner I. A.; Gabova A. V.; Chigaleychik L. A.; Karabanov A. V.* Investigation and Development of Methods for Automatic Search for AUC-diagram-based Features of Parkinson's Disease and Essential Tremor // RENSIT: Radioelectronics. Nanosystems. Information technologies. — М.: RANS, 2024. — Vol. 16, No. 1. — Pp. 67–78.

О методах машинного обучения в задаче gene promoter prediction

Дюкова Анастасия Петровна^{1*}

anastasia.d.95@gmail.com

Дюкова Елена Всеволодовна¹

edjukova@mail.ru

¹Москва, ФИЦ ИУ РАН

Геномы — символьные последовательности в алфавите $\{A, C, G, T\}$, буквы которого соответствуют нуклеотидам ДНК. Одним из основных этапов расшифровки геномных последовательностей является нахождение в них особых структурных единиц — генов. Принципиально важным является нахождение границ генов в последовательности генома и в связи с этим определение специальных участков генома (регуляторных участков экспрессии гена), называемых промоторами.

Для задачи распознавания промотора (англ.: gene promoter prediction) традиционно применяются алгоритмы машинного и глубокого обучения. По сути решается хорошо известная в интеллектуальном анализе данных задача классификации по прецедентам. В основном рассматриваются методы бинарной классификации, позволяющие отличать промоторный участок геномной последовательности модельного организма от непромоторного участка этого организма. Для формирования признаков, как правило, используются так называемые k -меры (упорядоченные наборы из k нуклеотидов). Частота встречаемости k -меры в рассматриваемой области генома задает значение соответствующего признака. Наилучшая точность классификации (порядка 97 %) получена методом опорных векторов [1] и с применением нейронной сети [2] на сбалансированных выборках.

Основной целью представляемой работы явилось исследование возможности применения логических методов анализа и классификации данных в задаче gene promoter prediction на примере организма Дрозофила фруктовой (*Drosophila melanogaster*). Эти методы базируются на поиске информативных фрагментов в признаковых описаниях прецедентов и ориентированы на обработку целочисленной информации низкой значности. Искомые фрагменты хорошо интерпретируемы и позволяют отличать промоторы от других областей генома, однако их поиск требует больших временных затрат [3, 4].

Для формирования выборки, содержащей 16000 примеров промоторных последовательностей организма и 42000 отрицательных примеров, представляющих собой участки экзонов, были использованы соответственно базы данных EPDNnew (<https://epd.expasy.org/epd>) и Flybase (<https://flybase.org>). В тестировании также участвовал ряд других алгоритмов машинного обучения: случайный лес, логистическая регрессия, LightGB, Catboost, XGBoost [5, 6].

Рассмотрен стандартный способ формирования признаков, использующий k -меры, и оригинальная методика прямого применения классификатора к исходным данным. Показано, что во втором случае качество логической классификации существенно выше и составляет 94,3 % по показателю ROC-AUC с ис-

пользованием ансамблевого подхода. Лучшие результаты, а именно, точность по ROC-AUC равную 95,1% и 94,8% показал классификатор Catboost соответственно при прямом применении к исходной выборке и с использованием k -мер. Исследованы вопросы выделения наиболее информативных k -мер.

- [1] Zhang M., Jia C., Li F., Li C., Zhu Y., Akutsu T., Webb G. I., Zou Q., Coin L. G. M., Song J. Critical Assessment of Computational Tools for Prokaryotic and Eukaryotic Promoter Prediction // Briefings in Bioinformatics. — 2022. — Vol. 23, No. 2. — bbab551.
- [2] Zhu Y., Li F., Xiang D., Akutsu T., Song J., Jia C. Computational Identification of Eukaryotic Promoters Based on Cascaded Deep Capsule Neural Networks // Briefings in Bioinformatics. — 2021. — Vol. 22, No. 4. — bbaa299.
- [3] Дюкова Е. В., Песков Н. В. Поиск информативных фрагментов описаний объектов в дискретных процедурах распознавания // Ж. вычисл. матем. и матем. физ. — 2002. — Т. 42., No 5. — С. 741–753.
- [4] Dragunov N., Djukova E., Djukova A. Supervised Classification and Finding Frequent Elements in Data // 8th Conf. Information Technology and Nanotechnology Proceedings. — NJ: IEEE, 2022. — Pp. 1–5.
- [5] Bishop C. M. Pattern Recognition and Machine Learning (Series: Information Science and Statistics). — Springer, 2006. — 738 p.
- [6] Friedman J. Stochastic Gradient Boosting // Computational Statistics & Data Analysis. — 2002. — Vol. 38. — Pp. 367–378.

Детекция и определение класса переломов ребер на основании данных компьютерной томографии с использованием методов глубокого обучения

Юрченко Александр Андреевич^{1,2*}

yurchenko.aa@phystech.edu

Матвеев Иван Алексеевич¹

matveev@frccsc.ru

¹Москва, Федеральный исследовательский центр «Информатика и управление» РАН

²Москва, Московский физико-технический институт (НИУ)

Задача обнаружения, локализации и классификации переломов ребер на основании данных компьютерной томографии является актуальной и значимой в медицинской диагностике и планировании лечения. Однако идентификация переломов ребер на тонкосрезовых КТ-изображениях с аксиальным пространственным разрешением 0,5–1,5 мм является трудоемкой задачей для медиков.

В данной работе рассматривается задача детектирования и классификации переломов ребер на основании изображений компьютерной томографии области грудной клетки человека (КТ ОГК) с использованием методов компьютерного зрения. Производится построение клинически применимой автоматической системы поддержки принятия решений при обнаружении и определении категорий переломов ребер по данным снимков КТ ОГК.

Предлагается использование двухэтапного алгоритма, на первом шаге которого осуществляется детектирование переломов на снимке с применением модели SEResNeXt и модулем комбинирования признаков описаний различных пространственных разрешений ViFPN. На втором — уменьшение доли ложноположительных предсказаний и классификация переломов на переломы с диастазом, без диастаза и консолидированные. Применяется подход к обучению модели детекции переломов ребер с использованием метода одновременного решения нескольких задач, одной из которых является задача семантической сегментации костных структур грудной клетки. Исследуется использование дополнительного сегментационного выхода модели детекции с механизмом внимания AttentionUNet [4].

Экспериментальное определение оптимальных архитектур моделей производится с использованием общедоступного набора данных RibFrac [1, 2], содержащего КТ-исследования пациентов с размеченными сегментационными масками переломов ребер. В данной работе предлагается изменение формализации постановки с задачи сегментации, которую принято рассматривать авторами работ по аналогичной тематике, на задачу детекции в ерминах ограничивающих параллелепипедов.

На основании результатов экспериментов демонстрируется, что предложенная система обеспечивает высокое качество обнаружения переломов ребер на наборе данных RibFrac с максимальным значением чувствительности обнаружения 0,935 при среднем количестве ложноположительных предсказаний на одно изображение — 4,7. Приводится сравнение полученного алгоритма с существующими методами решения задачи сегментации переломов ребер [1, 3] на основании количественных мер чувствительности и FROC.

Работа поддержана грантом РФФИ № 24-21-00330.

- [1] *Jin L., Yang J., Kuang K. et al.* Deep-Learning-Assisted Detection and Segmentation of Rib Fractures from CT Scans: Development and Validation of FracNet // *eBioMedicine*. — 2020. — Vol. 208. — P. 109919.
- [2] *Yang J., Shi R., Jin L. et al.* Deep Rib Fracture Instance Segmentation and Classification from CT on the RibFrac Challenge // *arXiv*. — 2024. — <https://arxiv.org/abs/2402.09372>
- [3] *Cao Z., Xu L., Chen D. et al.* A Robust Shape-Aware Rib Fracture Detection and Segmentation Framework With Contrastive Learning // *IEEE Transactions on Multimedia*. — 2023. — Vol. 25. — Pp. 1584–1591.
- [4] *Oktay O., Schlemper J., Folgoc L. L. et al.* Attention U-Net: Learning Where to Look for the Pancreas // *Medical Imaging with Deep Learning (MIDL)*, 2018, Amsterdam, The Netherlands. — 2018.

Задача обнаружения неточных повторов в структурах макромолекул

*Панкратов Антон Николаевич*¹*

pan@impb.ru

*Панкратова Наталья Михайловна*¹

pnm@impb.ru

¹Пушино, Институт математических проблем биологии РАН — филиал Федерального государственного учреждения «Федеральный исследовательский центр Институт прикладной математики им. М. В. Келдыша Российской академии наук»

Рассматривается проблема обнаружения неточных повторов как в первичных, так и в пространственных структурах белковых макромолекул. Для обнаружения повторов предложен универсальный спектральный метод, основанный на отклонении аппроксимаций сравниваемых сигналов, отображении результатов и их анализе на точечной матрице. Для аппроксимации символьных последовательностей предложено разложение их в пучок плавно меняющихся сигналов. Обнаружено, что такой подход, не использующий предопределенные и обученные на биоинформационном материале статистические модели, оказывается тем не менее весьма чувствительным инструментом обнаружения и изучения структуры повторов на уровне первичной последовательности.

Задача обнаружения повторов в пространственной структуре белка начинается с построения модели, за основу которой принимается пространственная ломаная. Следующим шагом является перевод координат точек ломаной в инвариантное по отношению к выбору системы координат представление. Например, можно построить аппроксимирующую кривую в пространстве, обладающую третьим порядком гладкости, что позволяет построить натуральное уравнение кривой, инвариантное к выбору системы координат. При этом задачей конструктивной аппроксимации является построение модели, грубой по отношению к пластическим деформациям макромолекулы, обусловленным тепловым движением атомов, экспериментальными погрешностями и результатами молекулярно-динамического моделирования. Обнаружено, что ортогональные разложения, обеспечивающие среднеквадратичное приближение, не способны обеспечить требуемый уровень грубости модели. Решением является использование суммирования ортогональных рядов методом Фейера или многочленов равномерного приближения Бернштейна высокого порядка.

При анализе полученной точечной матрицы, отображающей повторы в последовательности или структуре, полезным оказываются построения транзитивного замыкания и плотности корреляционного интеграла в зависимости от порога решающего правила.

- [1] *Pankratov A. N., Pankratova N. M.* Spectral Method for Detecting Inexact Repeats in Character Sequences // Pattern Recognition and Image Analysis. — 2022. — Vol. 32, No. 3. — Pp. 622–625.

Авторский указатель — Author Index

	А		Ж
Абакумов А. В.	62	Жарова М. А.	35
Авдеева З. К.	39	Жизняков А. Л.	49
Аверкин А. Н.	71		З
Ангальт Е. М.	15	Забежайло М. И.	88
Анциперов В. Е.	11	Закиров М. Н.	76
Арзангулян М. Э.	95	Захаров А. А.	49
	Б	Захарова М. В.	49
Белов А. С.	96	Зотов А. В.	50, 94
	В	Зырянов К. С.	65
Виноградов А. П.	15		К
Волков Е. Н.	71	Карабанов А. В.	98
Воробьев Т. К.	51, 94	Караваева К. Э.	77
Воронцов К. В.	47, 85	Карацуба Е. А.	59
	Г	Карканица А. В.	21
Габова А. В.	98	Картавцев А. А.	30
Гаврилов М. С.	82	Качанов В. В.	83
Генрихов И. Е.	23	Кершнер В. А.	11
Гнеушев А. Н.	53	Кершнер И. А.	98
Голубцов П. В.	42, 77	Князь В. В.	96
Греков М. М.	44	Комаров Д. В.	50, 94
Грушо А. А.	88	Копылов А. В.	28, 72
	Д	Котов А. Е.	16
Двоенко С. Д.	19	Краснопрошин В. В.	21
Дейцева А. Г.	13, 90	Красоткин С. А.	56
Дмитриенко А. Е.	53	Куличков С. Н.	76
Докукин А. А.	27, 92	Курбаков М. Ю.	72
Дулин С. К.	33		Л
Дюкова А. П.	25, 99	Ланге М. М.	18
Дюкова Е. В.	23, 25, 99	Лебедев М. А.	50, 91
	Е	Лемтюжникова Д. В.	39
Емельянов Г. М.	81		М
Еремеев С. В.	62, 64	Майсурадзе А. И.	79
Ермаков П. Г.	51, 91	Макеев Д. А.	77
		Марков С. И.	83
		Мартыненко А. С.	87

- Матвеев И. А. 16, 45, 52, 100
 Мироманов Д. Б. 87
 Михайлов Д. В. 81
 Морозов А. А. 98
 Мошканцев П. В. 96
 Мурашов Д. М. 61
 Мурашова Е. Д. 61
 Мурынин А. Б. 74
- Н**
- Неделько В. М. 20
 Новиков Д. И. 96
- О**
- Орешкин А. И. 65
- П**
- Панкратов А. Н. 102
 Панкратов Д. А. 64
 Панкратова Н. М. 102
 Панов А. С. 52
 Петрова П. А. 83
 Писковский В. О. 88
 Попов Д. Н. 79
- Р**
- Рихтер А. А. 74
 Родченко В. Г. 21
 Рудикова-Фронхёфер Л. В. .. 13, 90
 Рябцев А. Б. 33
- С**
- Семенов М. Е. 87
 Семёнова С. Э. 42
 Семенчук Н. В. 13, 90
 Сенько О. В. 27
 Середин О. С. 28, 72
 Скачков Н. А. 47
 Снитковский М. В. 68
 Соколов В. А. 45
 Сулимова В. В. 72
 Сушкова О. С. 98
 Сычугов А. А. 44
- Т**
- Тепляков Л. М. 45
 Торшин И. Ю. 24
- Ф**
- Фаломкин И. И. 68
 Фурс Е. В. 51, 91, 94
- Х**
- Харинов М. В. 67
 Хохлова М. Н. 98
 Хубиев К. Ю. 87
- Ц**
- Цурков В. И. 35
- Ч**
- Чигалейчик Л. А. 98
 Чуличков А. И. 12, 65, 68, 76
- Ш**
- Шапкина Н. Е. 42, 77
 Шарафиев А. Ф. 39, 82
 Шубочкин Е. И. 38
 Шушко Н. И. 57
- Ю**
- Юрченко А. А. 100
- А**
- Abdrakhmanova V. 32
- В**
- Varashov E. 55
- С**
- Chen Y. 31
- М**
- Matveev A. 70
 Matveev I. 70
 Maysuradze A. 46
- С**
- Sidorov L. 46

Song P.	37
Sorokin K.	32
	W
Wang W.	31

Содержание — Table of Contents

Интеллектуальный анализ данных / Data Mining	11
<i>Анциперов В. Е., Кершнер В. А.</i>	
Нейроморфное кодирование/декодирование потоков данных на основе модели точечных процессов Пуассона	11
<i>Чуличков А. И.</i>	
Анализ и интерпретация данных как технология получения знаний о реальности	12
<i>Семенчук Н. В., Рудикова-Фронхёфер Л. В., Дейцева А. Г.</i>	
Анализ социального настроения городского населения с помощью оценок спектральных плотностей с заданной точностью	13
<i>Ангальт Е. М., Виноградов А. П.</i>	
Ослабленные обобщенные прецеденты и критерии отбора вторичных кластеров	15
<i>Котов А. Е., Матвеев И. А.</i>	
Детекция аномалий в рядах событий при условии высоких требований ко времени обработки	16
Машинное обучение / Machine Learning	18
<i>Ланге М. М.</i>	
О теоретико-информационных границах точности в задачах кодирования источников сообщений и анализа данных	18
<i>Двоенко С. Д.</i>	
Кластеризация парных сравнений	19
<i>Неделько В. М.</i>	
Дистилляция знаний в применении к моделям бустинга	20
<i>Краснопрошин В. В., Карканица А. В., Родченко В. Г.</i>	
Система распознавания на основе общности свойств	21
<i>Генрихов И. Е., Дюкова Е. В.</i>	
Построение и исследование логических корректоров над произведением частичных порядков	23
<i>Торшин И. Ю.</i>	
Методы топологического анализа для порождения более информативных синтетических признаков на основе опорных цепей и произвольных метрических функций расстояния	24
<i>Дюкова Е. В., Дюкова А. П.</i>	
О логической классификации частично упорядоченных целочисленных данных	25

<i>Сенько О. В., Докукин А. А.</i>	
Новые методы построения оптимальных ансамблей	27
<i>Копылов А. В., Середин О. С.</i>	
Гармонические меры качества для многоклассовой классификации . .	28
<i>Картавцев А. А.</i>	
Повышение точности прогнозов в машинном обучении с использованием двухуровневого ансамбля решающих деревьев	30
Прогнозирование и рекомендательные системы / Forecasting and Recommendation Systems	31
<i>Chen Y., Wang W.</i>	
Generating Human Object Interaction via Neural Ordinary Differential Equations for Zero-shot Robot Manipulation	31
<i>Abdrakhmanova V., Sorokin K.</i>	
Methods for Enhancing Intelligent Type Inference for the Python Language	32
<i>Рябцев А. Б., Дулин С. К.</i>	
Повышение структурной согласованности в задаче поиска групп идентичных объектов	33
<i>Жарова М. А., Цурков В. И.</i>	
Применение бустинга в рекомендательных системах	35
<i>Song P.</i>	
A Probabilistic Framework of Shared Decision-Making	37
<i>Шубочкин Е. И.</i>	
Новые методы построения рекомендательных систем на основе данных о покупках	38
<i>Шарафиев А. Ф., Лемтюжникова Д. В., Авдеева З. К.</i>	
Анализ графов целей и задач стратегического планирования	39
Аналитика больших данных / Big Data Analytics	42
<i>Семёнова С. Э., Голубцов П. В., Шапкина Н. Е.</i>	
Анализ больших временных рядов с пропусками в реальном времени .	42
Нейронные сети и глубокое обучение / Neural Networks and Deep Learning	44
<i>Греков М. М., Сычугов А. А.</i>	
Модель автоматизированной системы имитации атак на основе машинного обучения	44
<i>Соколов В. А., Матвеев И. А., Тепляков Л. М.</i>	
Выделение эмбедингов для активного обучения, соответствующих регионам интереса на изображении	45

<i>Sidorov L., Maysuradze A.</i> The Influence of Graph Construction Methods on the Generalization Ability of Graph Neural Networks in Neurophysiological EEG Analysis	46
<i>Скачков Н. А., Воронцов К. В.</i> Метод маскировки входа для обучения моделей перевода	47
<i>Захаров А. А., Захарова М. В., Жизняков А. Л.</i> Обнаружение объектов на изображениях с использованием сверточных сетей на графах и ограниченного количества тренировочных данных	49
<i>Зотов А. В., Лебедев М. А., Комаров Д. В.</i> Нейросетевая детекция линий разметки взлетно-посадочной полосы и рулежных дорожек летного поля	50
<i>Воробьев Т. К., Ермаков П. Г., Фурс Е. В.</i> Нейросетевая детекция взлетно-посадочной полосы на изображениях от системы технического зрения воздушных судов гражданской авиации	51
<i>Панов А. С., Матвеев И. А.</i> Метод объединения специализированных нейросетей в множество многозадачных моделей путем выявления взаимосвязей между их внутренними представлениями	52
<i>Дмитриенко А. Е., Гнеушев А. Н.</i> Обучение устойчивой к геометрическим искажениям входных данных нейросетевой модели классификации изображений	53
Методы оптимизации для интеллектуального анализа данных / Data Mining Optimization Techniques	55
<i>Varashov E.</i> Cluster Merging Techniques for the Travelling Salesman Problem	55
<i>Красоткин С. А.</i> Обобщение полиномиального случая выпуклой оболочки для задачи коммивояжера: метод попарных сравнений	56
<i>Шушко Н. И.</i> Метод попарного сравнения для задачи двухуровневого голосования	57
Вычислительная сложность и приближенные методы / Algorithmic Complexity and Approximate Methods	59
<i>Карацуба Е. А.</i> О «сжатых» степенных рядах и сложности их вычисления	59
Обработка и анализ изображений, компьютерное зрение / Image Processing, Computer Vision	61

<i>Мурашов Д. М., Мурашова Е. Д.</i>	
Оценивание качества сегментации цифровых изображений	61
<i>Абакумов А. В., Еремеев С. В.</i>	
Оптимизация декомпозиции изображений по топологическим признакам с использованием персистентных ландшафтов	62
<i>Панкратов Д. А., Еремеев С. В.</i>	
Применение топологической декомпозиции изображений для классификации криогелей поливинилового спирта	64
<i>Зырянов К. С., Орешкин А. И., Чуличков А. И.</i>	
Интегрированный метод сегментации и уточнения рельефа атомных изображений на основе свертки формы зонда и нейронной сети архитектуры U-net	65
<i>Харинов М. В.</i>	
Генерация оптимальных приближений серого изображения	67
<i>Снитковский М. В., Чуличков А. И., Фаломкин И. И.</i>	
О задаче восстановления функциональной зависимости при наличии интенсивной импульсной помехи	68
<i>Matveev A., Matveev I.</i>	
Improving Text Detection by Predicting Missed Bounding Boxes Using Graph Models	70
<i>Волков Е. Н., Аверкин А. Н.</i>	
Генерация текстового описания ОКТ-снимков на русском языке с использованием визуально-языковых моделей	71
<i>Курбаков М. Ю., Сулимова В. В., Середин О. С., Копылов А. В.</i>	
Формирование интерпретируемых признаков на основе минимального остовного леса и их применение для определения структурированности расположения наночастиц на снимках с электронного микроскопа	72
<i>Рихтер А. А., Мурынин А. Б.</i>	
Алгоритмы морфологического анализа векторизованных границ на изображениях	74
Обработка и анализ сигналов / Signal Processing	76
<i>Закиров М. Н., Чуличков А. И., Куличков С. Н.</i>	
Декомпозиция сложных акустических атмосферных сигналов на части определенной формы	76
<i>Карабаева К. Э., Макеев Д. А., Шапкина Н. Е., Голубцов П. В.</i>	
Сравнение алгоритмов поиска аномалий, основанных на методах машинного обучения и спектрального анализа	77

<i>Майсурадзе А. И., Попов Д. Н.</i> Контролируемая генерация синтетических данных для задачи диаризации	79
Информационный поиск и анализ текстов / Information Retrieval and Text Analysis	81
<i>Михайлов Д. В., Емельянов Г. М.</i> Аннотации, вводные и заключительные разделы научных статей и их ранжирование по близости смысловому эталону	81
<i>Шарафиев А. Ф., Гаврилов М. С.</i> Использование графов знаний для решения задачи тематической сегментации научных текстов	82
<i>Петрова П. А., Марков С. И., Качанов В. В.</i> Создание набора данных для комбинированной классификации рецензий исходного кода	83
<i>Воронцов К. В.</i> Мастерская знаний: большие языковые модели для поиска и систематизации научной информации	85
Индустриальные приложения науки о данных / Industrial Data Science Applications	87
<i>Мартыненко А. С., Мироманов Д. Б., Семенов М. Е., Хубиев К. Ю.</i> Применение моделей детекции объектов для анализа видеопотоков в реальном времени	87
<i>Писковский В. О., Грушо А. А., Забезжайло М. И.</i> Задачи оптимизации использования распределенной облачной вычислительной инфраструктуры	88
<i>Рудикова-Фронхёфер Л. В., Семенчук Н. В., Дейцева А. Г.</i> О концепции построения универсальной системы городской среды . . .	90
<i>Ермаков П. Г., Лебедев М. А., Фурс Е. В.</i> Визуальная навигация беспилотного летательного аппарата на основе анализа изображений подстилающей поверхности	91
<i>Докукин А. А.</i> О построении системы считывания информации с экрана контроллера	92
<i>Фурс Е. В., Воробьев Т. К., Зотов А. В., Комаров Д. В.</i> Обеспечение ситуационной осведомленности экипажа сверхзвуковых гражданских самолетов	94
<i>Арзангулян М. Э.</i> Программный комплекс для анализа и прогнозирования временных рядов метеоданных	95

<i>Мошканцев П. В., Князь В. В., Новиков Д. И., Белов А. С.</i> Сегментация и нормализация семантической разметки крыш зданий по аэрофотосъемке для реконструкции трехмерных моделей районов жилой застройки	96
Анализ биомедицинских данных, биоинформатика / Analysis of Biomedical Data, Bioinformatics	98
<i>Сушкова О. С., Морозов А. А., Хохлова М. Н., Керинер И. А., Габова А. В., Чигалейчик Л. А., Карabanов А. В.</i> Обобщение AUC-диаграмм для автоматического поиска признаков болезни Паркинсона и эссенциального тремора	98
<i>Дюкова А. П., Дюкова Е. В.</i> О методах машинного обучения в задаче gene promoter prediction . . .	99
<i>Юрченко А. А., Матвеев И. А.</i> Детекция и определение класса переломов ребер на основании данных компьютерной томографии с использованием методов глубокого обучения	100
<i>Панкратов А. Н., Панкратова Н. М.</i> Задача обнаружения неточных повторов в структурах макромолекул .	102
Авторский указатель — Author Index	103
Содержание — Table of Contents	106

Научное издание

ИНТЕЛЛЕКТУАЛИЗАЦИЯ
ОБРАБОТКИ ИНФОРМАЦИИ

Тезисы докладов
15-й международной конференции
Гродно, 23–27 сентября, 2024 г.

Издаётся в авторской редакции
Ответственный за выпуск *А. Ф. Проневич*

Подписано в печать 16.09.2024. Формат 70×100¹/₁₆.
Бумага офсетная. Ризография. Гарнитура Computer Modern.
Усл.-печ. л. 9,1. Уч.-изд. л. 8,0. Тираж 100 экз. Заказ 058.

Издатель и полиграфическое исполнение:
учреждение образования
«Гродненский государственный университет
имени Янки Купалы».
Свидетельство о государственной регистрации издателя,
изготовителя, распространителя печатных изданий
№ 1/261 от 20.02.2024.
Ул. Ожешко, д. 22, 230023, Гродно.

ISBN 978-985-582-642-3



9 789855 826423