



Российская академия наук
Вычислительный центр РАН ФИЦ ИУ РАН
Центр хранения и анализа больших данных МГУ
Центр компетенций НТИ по направлению
«Искусственный интеллект» МФТИ

Интеллектуализация обработки информации

13-я Международная конференция

Москва, 2020

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

ББК 22.1:32.973.26-018.2

И 73

Интеллектуализация обработки информации: Тезисы докладов 13-й Международной конференции, г. Москва 2020 г. — М.: Российская академия наук, 2020. — 472 с.

ISBN 978-5-907366-16-9

В сборнике представлены тезисы докладов 13-й Международной конференции «Интеллектуализация обработки информации», проводимой Российской академией наук, Вычислительным центром Федерального исследовательского центра «Информатика и управление» РАН, Центром компетенций НТИ по направлению «Искусственный интеллект» на базе Московского физико-технического института, Центром хранения и анализа больших данных на базе Московского государственного университета имени М. В. Ломоносова.

Конференция проводится регулярно, начиная с 1989 г., и является представительным научным форумом в области интеллектуального анализа данных, машинного обучения, распознавания образов, анализа изображений, обработки сигналов, дискретного анализа.

Сайт конференции <http://mmro.ru>.

ISBN 978-5-907366-16-9

© Авторы докладов, 2020

© ФИЦ ИУ РАН, 2020

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

Intelligent Data Processing: Theory and Applications: Book of abstract of the 13th International Conference, Moscow, 2020. — Moscow: Russian Academy of Sciences, 2020. — 472 p.

ISBN 978-5-907366-16-9

The volume contains the abstracts of the 13th International Conference “Intelligent Data Processing: Theory and Applications”. The conference is organized by the Russian Academy of Sciences, Federal Research Center “Computer Science and Control” of RAS, Center of big data storage and analysis at the Moscow State University, and the Competence Center of the National Technological Initiative “Artificial intelligence” at the Moscow Institute of Physics and Technology.

The conference has being held biennially 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 conference website <http://mmro.ru/en/>.

ISBN 978-5-907366-16-9

© Authors of the abstracts, 2020
© FRC CSC RAS, 2020

Возможности использования деревьев решений в задаче атрибуции публицистических текстов XIX века

*Рогов Александр Александрович*¹

rogov@petrsu.ru

Москин Николай Дмитриевич^{1*}

moskin@petrsu.ru

*Абрамов Роман Владимирович*²

monset008@gmail.com

*Кулаков Кирилл Александрович*¹

kulakov@cs.karelia.ru

¹Петрозаводск, Петрозаводский государственный университет

²Санкт-Петербург, Национальный исследовательский университет ИТМО

В статье рассматривается математическая и программная поддержка задачи атрибуции (установления авторства) анонимных текстов [1]. Исследование проводилось на публицистических статьях XIX века из журналов «Время» (1861-1863), «Эпоха» (1864-1865) и еженедельника «Гражданин» (1873-1874). Известно, что Ф. М. Достоевский (вместе со своим братом М. М. Достоевским) редактировал и возглавлял эти журналы, поэтому уже давно ведутся исследования на предмет принадлежности его перу данных произведений. Большое количество этих статей опубликовано анонимно, т. е. либо без подписи, либо под псевдонимами. Впрочем, это относится и к статьям, которые исследователи давно приписывали Достоевскому, более или менее основываясь на документальных данных. Текст изучают на разных уровнях: пунктуационном, орфографическом, синтаксическом, лексико-фразеологическом и стилистическом (отметим, что под «авторским стилем» обычно понимают последние три уровня). Исследование проводилось с помощью информационной системы «Смалт» («Статистические методы анализа литературных текстов»), разработанной в Петрозаводском государственном университете [2].

В настоящее время для решения задачи атрибуции используют методы машинного обучения. В статье основное внимание уделено поиску признаков атрибуции с использованием технологий «дерева решений» и «случайный лес». Как известно, одним из достоинств данных технологий является хорошая интерпретация полученных результатов, что является ключевым при их признании специалистами в области филологии. Был исследован метод, основанный на дереве решений, для классификации статей на два класса: «Ф. М. Достоевский» и «другие». В качестве признаков были взяты статистики n -грамм частей речи (последовательностей из n закодированных частей речи). Для определения частотных характеристик текстов применялась информационная система СМАЛТ, а для построения деревьев решений использовалась среда Python 3.6. Используя лишь одну последовательность, удалось достичь результата 89% точности классификации текстов. С помощью полученных результатов было проанализировано влияние выбора глубины дерева, длины n -граммы, размера текста и остальных параметров на конечную точность алгоритма. Кроме того, были выявлены наиболее информативные части речи и их комбинации для данной задачи.

Ф. М. Достоевский хорошо понимал важность влияния сильных позиций текста на читателя, поэтому мог уделять больше внимания внесению правок в начальные и в конечные абзацы текстов чужих статей. Поэтому решая вопросы, связанные с атрибуцией текстов в журналах «Время» и «Эпоха», в качестве одной из задач выделен специальный анализ данных элементов текста. Была рассмотрена совокупность статей Ф. М. Достоевского и других авторов (М. М. Достоевский, Н. Н. Страхов, А. А. Головачев, И. Н. Шилль, А. Григорьев, А. У. Порецкий, Я. П. Полонский), опубликованных в этих журналах в период 1861-1865 гг. В текстах были выделены фрагменты размером 500, 700 и 1000 слов. При этом для увеличения объема выборки использовался шаг для отсчета начала следующего фрагмента: 100, 200 слов и т. п. На основе частеречного распределения фрагментов текстов были построены деревья решений, в узлах которых находятся условия ветвления, основанные на частоте встречаемости той или иной n -граммы. Анализ сильных позиций данных текстов с помощью деревьев решений показывает возможность стилистической правки, которую вносил Ф. М. Достоевский в тексты изначальных авторов.

Лексический спектр является значимой характеристикой для решения задачи определения авторства текстов (например, его использовал Г. Хетсо при исследовании текстов Ф. М. Достоевского). Однако применение деревьев решений требует представления спектра в виде одного числа, которое адекватно отражало бы его структуру. Была рассмотрена аппроксимация лексических спектров (на уровне словаря и на уровне текста) гиперболическими и экспоненциальными кривыми, в результате чего получаются две характеристики для каждой кривой. На материале статей из дореволюционного журнала «Время» (1861-1863) показано, что коэффициенты гиперболической регрессии аппроксимируют данные гораздо лучше, чем коэффициенты экспоненциальной кривой. Построение спектров осуществлялось с помощью информационной системы СМАЛТ.

При решении задачи атрибуции текстов возникает проблема определения авторского стиля писателя, который создал меньшее количество текстов (как количественно, так и по общему объему слов) в сравнении с другими авторами из числа анализируемых. Были рассмотрены возможные варианты решения этой проблемы на примере определения стиля Аполлона Григорьева. В качестве метода построения ансамбля классификаторов в работе использовался бэггинг (bootstrap aggregating). В результате расчетов выяснилось, что относительная частота биграммы «частица-прилагательное» больше 6,5 является отличительной особенностью публицистического стиля Аполлона Григорьева. Также было проведено исследование статьи «Стихотворения А. С. Хомякова», которое подтверждает ранее сделанный вывод о том, что нет оснований считать ее принадлежащей Аполлону Григорьеву.

Применение других методов машинного обучения (рекуррентные сети и параллельные рекуррентные сети) показали результаты, сравнимые с деревьями ре-

шений. Эффективнее оказалась модель трансформера, однако она требует большого объема данных для обучения.

Работа поддержана грантом РФФИ № 18-012-90026.

- [1] *Абрамов Р. В.* Применение дерева решений для атрибуции текстов // Процессы управления и устойчивость, 2020. Т. 7. № 1. С. 183–187.
- [2] *Кулаков К. А., Рогов А. А., Москвин Н. Д.* Программная поддержка в решении задачи атрибуции текстов // Программная инженерия, 2019. Т. 10. № 5. С. 234–240.

Содержание

Интеллектуальный анализ данных	10
<i>Забезжайло М. И.</i>	
О наследуемости диагностических заключений при пополнении обучающей выборки новыми эмпирическими данными	10
Машинное обучение	16
<i>Стрижов В. В., Адуенко А. А., Бажтеев О. Ю., Исаченко Р. В., Грабовой А. В.</i>	
Выбор моделей и ансамблей	16
<i>Неделько В. М.</i>	
О качественном поведении компонент разложений критериев качества решающих функций	18
<i>Бакланова А. О., Дюкова Е. В., Масляков Г. О.</i>	
Исследование зависимости качества классификации от выбора частных порядков на множествах значений признаков	21
<i>Нейчев Р. Г., Шибанов И. А., Стрижов В. В.</i>	
Восстановление графов суперпозиций функций в задаче символьной регрессии	26
<i>Ларин А. О., Середин О. С., Копылов А. В.</i>	
Критерий одноклассовой классификации при наличии нетипичных объектов в обучающей выборке	28
<i>Исаченко Р. В., Стрижов В. В.</i>	
Снижение размерности в задаче декодирования временных рядов	31
<i>Ланге М. М., Парамонов С. В.</i>	
Нижняя граница и избыточность вероятности ошибки классификации	33
<i>Бериков В. Б.</i>	
Ансамблевый кластерный анализ с использованием разнородного трансферного обучения	39
<i>Макарова А. И., Курбаков М. Ю., Сулимова В. В.</i>	
Нелинейный метод средних решающих правил с умными подвыборками для решения больших двухклассовых задач SVM-классификации	41
<i>Луканин А. А., Рязанов В. В.</i>	
Прогнозирование на базе решения набора задач классификации с учителем и степеней принадлежности	46
<i>Масич И. С., Краева Е. М.</i>	
Максимальные логические закономерности для построения решающих правил распознавания	48

<i>Полякова А. С., Липинский Л. В., Семенкин Е. С.</i> Исследование методов сокращения опорной выборки при коллективном выводе с помощью нечетких логических систем	53
<i>Потанин М. С., Стрижов В. В.</i> Аддитивная регуляризация для выбора структуры сетей глубокого обучения	57
Аналитика больших данных	59
<i>Генрихов И. Е., Дюкова Е. В.</i> О поиске частых элементов в небинарных данных на основе технологии CUDA	59
<i>Грабовой А. В., Стрижов В. В.</i> Задача обучения с экспертом для построение интерпретируемых моделей машинного обучения	64
<i>Сенько О. В., Добролюбова О. А.</i> Анализ временных рядов с учетом критерия стационарности	66
<i>Адуенко А. А., Стрижов В. В.</i> Выбор мультимodelей в задачах классификации и фильтрация выбросов	70
<i>Кириллюк И. Л., Сенько О. В.</i> Особенности группировки панельных данных на примере показателей, характеризующих экономическое развитие российских регионов	72
<i>Журавлёв Ю. И., Рязанов В. В., Сенько О. В., Докукин А. А., Виноградов А. П., Нелюбина Е. А., Стефановский Д. В.</i> Подход к использованию содержательного контекста для построения и численной проверки гипотез о скрытых закономерностях в данных	76
Нейронные сети и глубокое обучение	80
<i>Бериков В. Б.</i> Спектральный ансамблевый кластерный анализ с использованием малоранговых представлений и нейросетевого автоэнкодера	80
<i>Ефимов Ю. С., Соломатин И. А., Одиноких Г. А.</i> Поиск границ радужной оболочки при помощи сверточных нейронных сетей	86
<i>Нарцев Д. Ю., Гнеушев А. Н.</i> Оптимизация регрессионных нейросетевых моделей прямой оценки параметров объектов на изображениях модифицированными методами Adam	88

<i>Самсонов Н. А., Гнеушев А. Н., Матвеев И. А.</i> Построение нейросетевого классификатора в пространстве дескрипторов преобразования Радона для эффективного детектирования пешеходов	90
<i>Ахмедова Ш. А., Становов В. В., Камия Ю.</i> Автоматическое проектирование интерпретируемых ансамблей на основе нечетких систем и нейронных сетей	92
<i>Ямаев А. В., Чукалина М. П., Николаев Д. В., Шешкус А. В., Чуличков А. И.</i> Легковесная Шумоподавляющая Фильтрующая Нейронная Сеть Для Алгоритма FBP	98
<i>Привезенцев Д. Г., Жизняков А. Л., Белякова А. С.</i> Анализ и прогнозирование уровня глюкозы в крови на основе нейронных сетей и данных суточного мониторинга	104
<i>Никитин Ф. А., Стрижов В. В.</i> Графовые нейронные сети для несвязанных графов в химических реакциях.	106
<i>Князь В. В., Мизгинов В. А., Гродзицкий Л. В., Мошканцев П. В.</i> Оценка качества работы нейронной сети для предсказания 3D модели объекта	108
<i>Гродзицкий Л. В., Данилов С. Ю., Князь В. В.</i> Глубокое обучение для задач отображения улучшенного видения на ИЛС112	
<i>Гогоберидзе Ю. Т., Классен В. И., Натензон М. Я., Просвиркин И. А., Сафин А. А.</i> Опыт применения многослойных свёрточных нейронных сетей и технологий Big Data на примере искусственного медицинского интеллекта ФтизисБиоМед	115
Методы оптимизации для интеллектуального анализа данных	121
<i>Шибзухов З. М.</i> Об одном робастном подходе к поиску центров кластеров	121
<i>Немирко А. П., Дула Х.</i> Алгоритм ближайших выпуклых оболочек на основе использования линейного программирования	123
<i>Сороковиков П. С., Горнов А. Ю.</i> Трехэтапная вычислительная технология оптимизации атомно-молекулярных кластеров Морса сверхбольших размерностей	125
<i>Аникин А. С.</i> Численное решение задач минимизации потенциала Китинга с размерностями до 300 миллионов переменных	129

<i>Ерохин В. И., Кадочников А. П., Сотников С. В., Маркина М. К.</i> Оптимизационные методы численного решения систем линейных интервальных уравнений, связанных с задачами построения линейных зависимостей при интервальной неопределенности данных	131
<i>Запрднюк Т. С., Горнов А. Ю.</i> Методика стресс-тестирования программных комплексов для оптимизации нелинейных управляемых динамических систем	137
<i>Гладин Е. Л.</i> Линейная сходимость в гладкой выпуклой задаче min-min с сильной выпуклостью по одной из групп переменных	139
<i>Тримбач Е. А., Рогозин А. В.</i> Ускорение стохастических методов на примере децентрализованного SGD	143
Вычислительная сложность и приближенные методы	145
<i>Казаковцев Л. А., Рожнов И. П., Попов А. М., Товбис Е. М.</i> Самонастраивающийся алгоритм поиска с чередующимися окрестностями для почти оптимального решения задачи кластеризации k-средних	145
<i>Карацуба Е. А.</i> Интервальный подход в проблеме аппроксимации Эйлеровой константы	150
<i>Ташилинский А. Г., Сафина Г. Л.</i> Вероятностное моделирование процесса стохастического оценивания межкадровых геометрических деформаций изображений	154
<i>Белозуб В. А., Козлова М. Г., Лукьяненко В. А.</i> Восстановление решений уравнений типа Урысона	158
Обработка и анализ изображений и сигналов, компьютерное зрение	162
<i>Арсеев С. П., Местецкий Л. М.</i> Скелет символа как модель следа пера для распознавания по восстановленной траектории	162
<i>Мурыннин А. Б., Матвеев И. А., Игнатьев В. Ю.</i> Применение генеративных нейросетей для повышения пространственного разрешения спутниковых изображений	167
<i>Липкина А. Л., Местецкий Л. М.</i> Метод распознавания шрифтов на основе медиального представления	169
<i>Василенко В. В., Сафронов А. П., Смыслов А. А., Цепляев Д. П., Марковский А. Н.</i> Бигармоническое сглаживание изображений	171
<i>Мурашов Д. М.</i> Информационная модель для метода обеспечения качества автоматической сегментации изображений	173

<i>Фурсов В. А., Минаев Е. Ю., Котов А. П.</i> Определение Движения Оптической Системы на Основе Метода Согласованного Оценивания	178
<i>Применко Д. В., Панищев В. С., Бурцев О. А.</i> Определение эффективности алгоритмов выделения линий на изображении	184
<i>Елизаров А. А., Разинков Е. В.</i> Метод повышения точности классификации изображений с использованием обучения с подкреплением	188
<i>Визильтер Ю. В., Выгоов О. В., Желтов С. Ю., Брянский С. А.</i> «Формула Эйлера» для морфологического анализа мозаичных изображений	191
<i>Ханьков И. Г., Ненашев В. А.</i> Применение квазиоптимальной кластеризации пикселей в задаче комплексования разноразмерных изображений	199
<i>Калмыков Н. С.</i> Обнаружение и сопровождение целей с БПЛА при помощи нейронных сетей	205
Информационный поиск и анализ текстов	208
<i>Воронцов К. В.</i> Десять открытых проблем вероятностного тематического моделирования	208
<i>Бахтеев О. Ю., Кузнецова Р. В., Хазов А. В., Огальцов А. В., Сафин К. Ф., Горленко Т. А., Суворова М. А., Иващенко А. А., Чехович Ю. В., Моттль В. В.</i> Поиск почти-дубликатов в рукописных текстах школьных сочинений	214
<i>Михайлов Д. В., Емельянов Г. М.</i> Оценка близости смысловому эталону без поиска перифраз и иерархия тематических текстов	219
<i>Евсеев Д. А., Архипов М. Ю.</i> Генерация SPARQL-запросов для ответа на сложные вопросы с помощью BERT и BiLSTM	225
<i>Елизаров А. М., Липачев Е. К.</i> Методы Big Math и интеграция математических знаний	227
<i>Кузнецова Р. В.</i> Вариационное моделирование правдоподобия с триплетными ограничениями в задачах информационного поиска	232
<i>Кальян В. П.</i> Математические и лингвистические аспекты моделирования медиадискурса	237

<i>Беленькая О. С., Суворова М. А., Филиппова О. А., Чехович Ю. В.</i> Задачи систем обнаружения заимствований в применении к поиску заимствований в учебных работах средней школы	239
Индустриальные приложения науки о данных	244
<i>Дедкова А. А., Флоринский И. В.</i> Анализ рельефа кремниевых пластин методами геоморфометрии	244
<i>Никулин В. С., Пестунов А. И.</i> Метод обеспечения отказоустойчивости вычислительных комплексов на основе оценки характеристик надежности	246
<i>Авдеева З. К., Гребенюк Е. А., Коврига С. В.</i> Мониторинг и прогнозирование в слабо-структурированных ситуациях с использованием временных рядов и когнитивного моделирования	250
<i>Андрянов Н. А., Дементьев В. Е., Ташилинский А. Г.</i> Применение многомерных моделей гауссовых смесей для анализа заказов службы такси	255
<i>Некрасов И. В.</i> Идентификация штатной работы оборудования на основе прямых геометрических методов	259
<i>Кульков Я. Ю., Жизняков А. Л., Привезенцев Д. Г., Запатрин М. Г.</i> Использование системы технического зрения в системе прослеживания производства железнодорожных колес	265
<i>Кульков Я. Ю., Садыков С. С., Орлов А. Д., Баюров С. В.</i> Вычисление координат точки захвата плоского объекта роботом	269
<i>Астафьев А. В., Демидов А. А., Кондрушин И. Е., Макаров М. В.</i> Разработка алгоритма позиционирования объекта по данным с активной сенсорной сети Bluetooth Low Energy маяков	273
<i>Старожилец В. О., Чехович Ю. В.</i> Об одном подходе к статистическому моделированию транспортных потоков на МКАД и управлению въездами	277
Анализ биомедицинских данных, биоинформатика	281
<i>Сушкова О. С., Морозов А. А., Габова А. В., Карабанов А. В.</i> Разработка метода ранней и дифференциальной диагностики болезни Паркинсона и эссенциального тремора с помощью анализа всплескообразной активности мышц	281
<i>Янковская А. Е., Обуховская В. Б.</i> Расширение прикладной интеллектуальной системы диагностики качества жизни пациентов с неврологической патологией с учетом психологической безопасности	283

<i>Кершнер И. А., Обухов Ю. В., Синкин М. В.</i>	
Сегментация длительных сигналов ЭЭГ на области интереса и способ дифференциации эпилептических приступов от артефактов жевания	287
<i>Рыкунов С. Д., Бойко А. И., Маслова О. А., Устинин М. Н.</i>	
Реконструкция функциональной структуры мозга человека по данным электроэнцефалографии	289
<i>Устинин М. Н., Рыкунов С. Д., Бойко А. И.</i>	
Реконструкция пространственной структуры нервной и мышечной системы тела человека по его магнитному полю	291
<i>Толмачева Р. А., Обухов Ю. В., Жаворонкова Л. А.</i>	
Мониторинг межканальной фазовой синхронизации ЭЭГ у пациентов с черепно-мозговой травмой до и после реабилитации	293
<i>Сенько О. В., Салманов М. Ю., Брусов О. С., Матвеев И. А., Кузнецова А. В.</i>	
Метод генерации признаков описаний, основанный на расстояниях до эталонов, в биомедицинских исследованиях	295
<i>Ройзензон Г. В., Соколов А. В., Черешкин Д. С., Комендантова Н. П., Голубков В. В., Бритков В. Б.</i>	
Пандемия Covid19 и методы интеллектуального анализа рисков	301
<i>Соколов А. В., Соколова Л. А.</i>	
Технология сбалансированной идентификации: выбор модели динамики COVID-19 по имеющимся данным	307
<i>Руднев В. Р., Куликова Л. И., Кайшева А. Л., Тихонов Д. А.</i>	
Разработка и развитие базы данных двухспиральных мотивов белковых молекул и вычислительные сервисы для их анализа	312
<i>Гончаренко В. В., Григорян Р. К., Самохина А. М.</i>	
Подходы к мультиклассовой классификации датасета потенциалов R300	315
<i>Куликов А. М., Харламов А. А.</i>	
Использование однородной семантической сети для классификации результатов генетического анализа	318
<i>Харламов А. А.</i>	
Один тип искусственной нейронной сети на основе нейронов с временной суммацией сигналов	320
Методы математического моделирования в интеллектуальном анализе данных	322
<i>Двоенко С. Д., Пшеничный Д. О.</i>	
О новых типах медианы Кемени	322
Интеллектуальный анализ геопространственных данных	324

<i>Мандрикова О. В., Родоманская А. И.</i> Вейвлет-модель вариаций геомагнитного поля	324
<i>Геппенер В. В., Мандрикова Б. С.</i> Метод обнаружения аномальных эффектов в сложном сигнале	326
<i>Кошелева Н. В., Гвоздев О. Г., Козуб В. А., Мурынин А. Б., Рихтер А. А.</i> Восстановление 3D-модели объектов инфраструктуры на основе использования нейросетевых методов обработки спутниковых изображений	330
<i>Филлин А. И., Грачева И. А., Копылов А. В.</i> Совместная оценка карты рассеивания и атмосферной освещенности с использованием вероятностной гамма-нормальной модели для задачи устранения тумана на изображении	332
<i>Рогов А. А., Москвин Н. Д., Абрамов Р. В., Кулаков К. А.</i> Возможности использования деревьев решений в задаче атрибуции публицистических текстов XIX века	336
Интеллектуальная оптимизация и эффективный менеджмент	341
<i>Становов В. В., Ахмедова Ш. А., Семенкин Е. С.</i> Комбинированный метод учета эpsilon-ограничений для решения задачи распределения нагрузки с помощью дифференциальной эволюции	341
<i>Семенкина О. Е., Попов Е. А., Семенкин Е. С.</i> Бионические алгоритмы для для оптимизации расписания в промышленности	346
<i>Германчук М. С., Козлова М. Г., Лукьяненко В. А.</i> Знаниеориентированные модели маршрутизации многих коммивояжеров	352
<i>Жукова Г. Н., Ульянов М. В.</i> Классификация асимметричных задач коммивояжера по квантилям распределения сложности индивидуальных задач	356
<i>Токарева В. А.</i> Эвристическая ребалансировка на основе приоритетов в задаче управления данными с вероятностными ограничениями	360
<i>Некрасов И. В., Правдивец Н. А.</i> Модели координации задач планирования закупки сырья и выпуска конечной продукции промышленного предприятия	362
<i>Федосенко Ю. С., Хандурин Д. К., Шеянов А. В.</i> Задача о биназначениях в приложении к проблеме воднотранспортного обслуживания островных и городских агломераций	368
<i>Куприянов Б. В., Лазарев А. А.</i> Решение задачи минимизации времени выполнения заказа для рекурсивного конвейера	374

Афраймович Л. Г., Емелин М. Д.

Стратегии комбинирования решений трехиндексной задачи о назначениях 378

Скобелев П. О., Ларюхин В. Б.

О проекте цифровой эко-системы для создания виртуального рынка цифровых двойников предприятий электротехнической промышленности 384

Джуманов Р. Р., Хуснуллин Н. Ф., Лазарев А. А.

Математическое моделирование планирования подготовки космонавтов 388

Макаровских Т. А., Панюкова А. А.

Распределение комплекса работ по исполнителям 390

Гришин Е. М.

Методы автоматической сборки белков 394

Гафаров Е. Р., Долгий А. Б., Сомов М. Л.

Верхние и нижние границы параллельного партийного планирования для одной машины с учетом последовательности работ 398

Галахов С. А.

Методы повышения эффективности энергетических систем 400

Сидельников Ю. В.

Расширение возможностей метрического подхода на основе теории средних и теории ошибок 402

Барашов Е. Б., Лазарев А. А., Правдивец Н. А.

Аппроксимация целевой функции задач теории расписаний 404

Лазарев А. А., Лемтюжникова Д. В.

Метрический подход для задач железнодорожного планирования . . . 410

Лемтюжникова Д. В., Тюняткин А. А.

Метрическая интерполяция в задачах теории расписаний 412

Мандель А. С., Лаптин В. А.

Оптимальное управление системами массового обслуживания в условиях применения для описания их состояния методов структурно-классификационной и экспертно-статистической обработки 414

Резников М. Б., Федосенко Ю. С.

О потенциале кластерных схем синтеза оптимальных расписаний для моделей воднотранспортной логистики 418

Саратов А. А.

Жадный алгоритм решения классической NP-трудной задачи минимизации суммарного запаздывания 423

Интеллектуальный анализ данных в задачах информационной безопасности . . 429

<i>Морозов А. А., Сушкова О. С.</i>	
Разработка общедоступного набора видеоданных в терагерцовом диапазоне и программной платформы для экспериментов с интеллектуальным видеонаблюдением в терагерцовом диапазоне	429
<i>Чигринский В. В., Матвеев И. А.</i>	
Оптимизация работы системы слежения, основанной на сети камер видеонаблюдения	431
<i>Григорьев А. Д., Гнеушев А. Н.</i>	
Слежение за множеством объектов на видео изображениях с помощью ре-идентификации с предфильтрацией дескрипторов по качеству . . .	433
<i>Сычугов А. А., Токарев В. Л.</i>	
Метод размещения иммунных детекторов на основе оценки риска безопасности сетевых узлов	435
Содержание	443
Авторский указатель	462

Contents

Data mining	10
<i>Zabekhailo M.</i>	
To the heritability of diagnostic conclusions at extension of training sample by new empirical data	13
Machine learning	16
<i>Strijov V., Aduenko A., Bakhteev O., Isachenko R., Grabovoy A.</i>	
Selection of models and ensembles	17
<i>Nedel'ko V.</i>	
On the Shape of Components of Decompositions of Quality Criteria for Decision Functions	20
<i>Baklanova A., Djukova E., Masliakov G.</i>	
Investigation of the dependence of the supervised classification quality on the choice of partial orders on feature values sets	24
<i>Neychev R., Shibayev I., Strijov V.</i>	
Optimal superposition trees restoration in symbolic regression	27
<i>Larin A., Seredin O., Kopylov A.</i>	
Criterion for one-class classification in the presence of outliers in the training set	30
<i>Isachenko R., Strijov V.</i>	
Dimensionality reduction for time series decoding	32
<i>Lange M., Paramonov S.</i>	
A lower bound and a redundancy of classification error probability	36
<i>Berikov V.</i>	
Ensemble Clustering with Heterogeneous Transfer Learning	40
<i>Makarova A., Kurbakov M., Sulimova V.</i>	
Smart Sample Kernel-based Mean Decision Rules Method for Big Binary SVM Classification Problems	44
<i>Lukanin A., Ryazanov V.</i>	
Prediction based on the solution of the set of classification problems of supervised learning and degrees of membership	47
<i>Masich I., Kraeva E.</i>	
Maximum logical patterns for constructing decision rules for recognition	51
<i>Polyakova A., Lipinskiy L., Semenkin E.</i>	
Investigation of Reference Sample Reduction Methods for Ensemble Output with Fuzzy Logic-Based Systems	55

<i>Potanin M., Strijov V.</i> Additive regularization schedule for neural architecture search	58
Big data analytics	59
<i>Genrikhov I., Djukova E.</i> About searching for frequent elements in nonbinary data based on CUDA technology	62
<i>Grabovoy A., Strijov V.</i> Expert learning for interpretable model selection	65
<i>Senko O., Dobroliubova O.</i> Time series analysis with stationarity criterion	68
<i>Aduenko A., Strijov V.</i> Multimodel Selection for Classification and Outlier Filtering	71
<i>Kirilyuk I., Senko O.</i> Peculiarities of grouping of panel data on the example of indicators character- ising the economic development of Russian regions	74
<i>Zhouravlev Yu., Ryazanov V., Senko O., Dokukin A., Vinogradov A., Nelyu- bina E., Stefanovskiy D.</i> An approach to using meaningful context to construct and numerically test hypotheses about hidden regularities in data	78
Neural networks and deep learning	80
<i>Berikov V.</i> Low-Rank Spectral Ensemble Clustering Using Autoencoder Network . . .	83
<i>Efimov Yu., Solomatin I., Odinokikh G.</i> Finding the borders of the iris using convolutional neural networks	87
<i>Nartsev D., Gneushev A.</i> Optimization of regression neural network models for direct estimation of object parameters in images by modified Adam methods	89
<i>Samsonov N., Gneushev A., Matveev I.</i> Neural network classifier in the space of Radon Transform descriptors for efficient pedestrian detection	91
<i>Akhmedova S., Stanovov V., Kamiya Y.</i> Automated design of interpretable ensembles based on fuzzy systems and neural networks	95
<i>Yamaev A., Chukalina M., Nikolaev D., Sheshkus A., Chulichkov A.</i> Lightweight Denoising Filtering Neural Network For FBP Algorithm. . . .	101

<i>Privezentsev D., Zhiznyakov A., Belyakova A.</i> Analysis and prediction of blood glucose levels based on neural networks and daily monitoring data	105
<i>Nikitin F., Strijov V.</i> Graph neural networks for disconnected graphs in chemical reactions. . . .	107
<i>Kniaz V., Mizginov V., Grodzitsky L., Moshkantsev P.</i> 3D Reconstruction Neural Network Quality Evaluation	110
<i>Grodzitsky L., Danilov S. Yu., Kniaz V.</i> Deep Learning for Projection of the Enhanced Vision on the HUD	114
<i>Gogoberidze Y., Klassen V., Natenzon M., Prosvirkin I., Safin A.</i> Experience of using multilayer convolutional neural networks and Big Data technologies on the example of artificial medical intelligence FtisisBioMed	118
Data mining optimization techniques	121
<i>Shibzukhov Z.</i> About one robust approach to the search for cluster centers	122
<i>Nemirko A., Dula J.</i> Nearest convex hulls algorithm based on linear programming — IDP-13 .	124
<i>Sorokovikov P., Gornov A.</i> Three-stage computational technology for optimization of atomic-molecular Morse clusters of extremely large dimensions	127
<i>Anikin A.</i> Numerical solution of Keating potential minimization problems with di- mensions up to 300 million variables	130
<i>Erokhin V., Kadochnikov A., Sotnikov S., Markina M.</i> Optimization methods for the numerical solution of systems of linear in- terval equations associated with the problems of constructing linear depen- dencies with interval data uncertainty	134
<i>Zarodnyuk T., Gornov A.</i> Stress testing technique of numerical investigating software for optimization of nonlinear controlled dynamical systems	138
<i>Gladin E.</i> Linear convergence for smooth convex min-min problem with strong con- vexity in one of the groups of variables	141
<i>Trimbach E., Rogozin A.</i> Acceleration of stochastic methods on the example of decentralized SGD .	144
Algorithmic complexity and approximate methods	145

<i>Kazakovtsev L., Rozhnov I., Popov A., Tovbis E.</i> Self-Adjusting Variable Neighborhood Search Algorithm for Near-Optimal k-Means Clustering	148
<i>Karatsuba E.</i> The interval approach in the problem of approximation of the Euler constant	152
<i>Tashlinskii A., Safina G.</i> Probabilistic modeling of stochastic estimation process of image inter-frame geometric deformations	156
<i>Belozb V., Kozlova M., Lukianenko V.</i> Reconstruction of solutions of equations of Uryson type	160
Image and signal processing, computer vision	162
<i>Arseev S., Mestetskiy L.</i> Symbol skeleton as a pen trace model for recognition using reconstructed trace	165
<i>Murynin A., Matveev I., Ignatiev V.</i> Application of generative neural networks to increase the spatial resolution of satellite images	168
<i>Lipkina A., Mestetskiy L.</i> Medial representation based font recognition method	170
<i>Vasilenko V., Safronov A., Smylov A., Tsyplyaev D., Markovskiy A.</i> Biharmonic smoothing the images — IDP-13	172
<i>Murashov D.</i> Information model for quality assessment method applied to automatic im- age segmentation	176
<i>Fursov V., Minaev E., Kotov A.</i> Motion Detection of Optical Systems Based on the Conformed Estimation Method	181
<i>Primenko D., Panishchev V., Burtsev O.</i> Determining the effectiveness of line identification algorithms in an image	186
<i>Elizarov A., Razinkov E.</i> Image classification accuracy improvement method using reinforcement learning	190
<i>Vizilter Yu., Vygolov O., Zheltov S., Brianskiy S.</i> “Euler Identity” for Morphological Image Analysis	195
<i>Khanykov I., Nenashev V.</i> The application of quasi-optimal pixel clustering in the problem of combin- ing multi-angle images	202

<i>Kalmykov N.</i>	
Target detection and tracking using neural networks on UAV	207
Information retrieval and text analysis	208
<i>Vorontsov K.</i>	
Ten open problems in probabilistic topic modeling	211
<i>Bakhteev O., Kuznetsova R., Khazov A., Ogaltsov A., Safin K., Gorlenko T., Suvorova M., Ivahnenko A., Chekhovich Y., Mottl V.</i>	
Near-duplicate detection in handwritten school essays	217
<i>Mikhaylov D., Emelyanov G.</i>	
Estimation for the closeness to a semantic pattern without paraphrasing, and a hierarchy of topical texts	222
<i>Evseev D., Arkhipov M.</i>	
SPARQL query generation for complex question answering with BERT and BiLSTM-based model	226
<i>Elizarov A., Lipachev E.</i>	
Big Math Methods and Mathematical Knowledge Integration	230
<i>Kuznetsova R.</i>	
Variational Bi-domain Triplet Modeling in Information Retrieval	235
<i>Kaliyan V.</i>	
Mathematical and linguistic aspects of media discourse modeling	238
<i>Belenkaya O., Suvorova M., Filippova O., Chekhovich Y.</i>	
Tasks of text reuse detection systems when applied to the text reuse detec- tion in secondary school written works	242
Industrial data science applications	244
<i>Dedkova A., Florinsky I.</i>	
Analysis of topography of silicon wafers by geomorphometric methods	245
<i>Nikulin V., Pestunov A.</i>	
Method of maintaining fault tolerance of computing systems based on the assessment of reliability characteristics	248
<i>Avdeeva Z., Grebenyuk E., Kovriga S.</i>	
Monitoring and forecasting in ill-structured situations based on time series and cognitive modelling	253
<i>Andriyanov N., Dementiev V., Tashlinskii A.</i>	
Application of multi-dimensional models of Gaussian mixtures models for analysis of taxi service	257
<i>Nekrasov I.</i>	
Direct Geometric Approach for Asset Normal State Identification	262

<i>Kulkov Y., Zhiznyakov A., Privezentsev D., Zapatrin A.</i> Use of a machine vision system for tracking the production of railway wheels	267
<i>Kulkov Y., Sadykov S., Orlov A., Bayurov S.</i> Calculation of the flat objects gripping point coordinates by a robot . . .	271
<i>Astafiev A., Demidov A., Kondrushin I., Makarov M.</i> Development of algorithm for positioning an object according to data from an active sensor network of Bluetooth Low Energy beacons	275
<i>Starozhilets V., Chekhovich Y.</i> About one approach to traffic flows statistical modeling on Moscow Ring Road and enters control	279
Analysis of biomedical data, bioinformatics	281
<i>Sushkova O., Morozov A., Gabova A., Karabanov A.</i> Development of a method for early and differential diagnosis of Parkinson's disease and essential tremor based on analysis of wave train electrical activity of muscles	282
<i>Yankovskaya A., Obukhovskaya V.</i> An expansion of applied intelligent system for diagnosing the quality of life of patients with neurological pathology with considering psychological safety	285
<i>Kershner I., Obukhov Yu., Sinkin M.</i> Segmentation of long-term EEG signals on the area of interest and a method for differentiating epileptic seizures from chewing artifacts	288
<i>Rykunov S., Boyko A., Maslova O., Ustinin M.</i> Reconstruction of the Human Brain Functional Structure Based on the Electroencephalography Data	290
<i>Ustinin M., Rykunov S., Boyko A.</i> Reconstruction of the spatial structure of the human body nervous and muscular systems based on its magnetic field	292
<i>Tolmacheva R., Obukhov Yu., Zhavoronkova L.</i> Monitoring of inter-channel EEG phase synchronization in patients with traumatic brain injury before and after rehabilitation	294
<i>Senko O., Salmanov M., Brusov O., Matveev I., Kuznetsova A.</i> The feature descriptions generating method based on distances to standards in biomedical research	298
<i>Royzenzon G., Sokolov A., Chereshkin D., Komendantova N., Golubkov V., Britkov V.</i> Covid19 pandemic and artificial intelligence methods for risk analysis . . .	304
<i>Sokolov A., Sokolova L.</i> Balanced Identification Technology: Choosing COVID-19 Dynamics Model for Available Data	310

<i>Rudnev V., Kulikova L., Kaysheva A., Tikhonov D.</i> Creation and development of a database of two helical motifs of protein molecules and computational services for their analysis	314
<i>Goncharenko V., Grigoryan R., Samokhina A.</i> Approaches to multiclass classification of the P300 dataset	317
<i>Kulikov A., Kharlamov A.</i> Using a homogeneous semantic network to classify the result of genetic analysis	319
<i>Kharlamov A.</i> On a type of artificial neural network based on neurons with temporal summation of signals	321
Methods of mathematical modeling in data mining	322
<i>Dvoenko S., Pshenichny D.</i> On New Types of the Kemeny's Median	323
Geospatial data mining	324
<i>Mandrikova O., Rodomanskay A.</i> Wavelet model of geomagnetic field variations	325
<i>Geppener V., Mandrikova B.</i> Method for detecting anomalous effects in a complex signal	328
<i>Kosheleva N., Gvozdev O., Kozub V., Murynin A., Richter A.</i> Reconstruction of a 3D model of infrastructure objects based on the usage of neural network methods for processing satellite images	331
<i>Filin A., Gracheva I., Kopylov A.</i> Combined transmission map estimation and atmospheric-light extraction using the probabilistic gamma-normal model for haze removal problem . .	334
<i>Rogov A., Moskin N., Abramov R., Kulakov K.</i> Possibilities of using decision trees in the problem of attribution of publicistic texts of the XIX century	339
Intelligent optimization and effective management	341
<i>Stanovov V., Akhmedova S., Semenkin E.</i> Combined Epsilon-Constraint Handling Method for Solving Economic Load Dispatch Problems with Differential Evolution	344
<i>Semenkina O., Popov E., Semenkin E.</i> Nature-inspired algorithms for scheduling optimization in industry	349
<i>Germanchuk M., Kozlova M., Lukianenko V.</i> Knowledgeoriented routing models for many traveling salesmen	354

<i>Zhukova G., Ulyanov M.</i>	
Classification of asymmetric traveling salesman problems by quantiles of the distribution of the complexity of individual problems	358
<i>Tokareva V.</i>	
Priority-based rebalancing heuristic for a mixed shop problem with probabilistic constraints	361
<i>Nekrasov I., Pravdivets N.</i>	
Coordination Models for Purchasing and Production Scheduling Processes of an Industrial Enterprise	365
<i>Fedosenko Y., Khandurin D., Sheyanov A.</i>	
Bi-assignment problem application to the problem of water transport services for island and urban agglomerations	371
<i>Kupriyanov B., Lazarev A.</i>	
Solving the problem of minimizing order lead time for a recursive conveyor	376
<i>Afraimovich L., Emelin M.</i>	
Strategies for combining solutions to a three-index assignment problem . .	381
<i>Skobelev P., Laryukhin V.</i>	
About the digital ecosystem project to create a virtual market for digital twins of electrical industry enterprises	386
<i>Jumanov R., Husnullin N., Lazarev A.</i>	
Mathematical modeling of cosmonaut training planning	389
<i>Makarovskikh T., Panyukova A.</i>	
Distribution of the Complex of Jobs by Performer	392
<i>Grishin E.</i>	
Methods of automated protein assembly	396
<i>Gafarov E., Dolgui A., Somov M.</i>	
On lower and upper bounds for single machine parallel batch scheduling subject to chains of jobs.	399
<i>Galakhov S.</i>	
Methods for improving the efficiency of energy systems	401
<i>Sidelnikov Yu.</i>	
Expansion of the possibilities of metric approach on the basic of the theory of averages and theory of the errors	403
<i>Barashov E., Lazarev A., Pravdivets N.</i>	
Approximation of the objective function of scheduling problems	407
<i>Lazarev A., Lemtyuzhnikova D.</i>	
Metric approach for railway planning problems	411
<i>Lemtyuzhnikova D., Tyunyatkin A.</i>	
Metric interpolation in scheduling problems	413

<i>Mandel A., Laptin V.</i>	
Optimal control of queuing systems (QS) when using the methods of structural-classification and expert-statistical processing to describe the QS state	416
<i>Reznikov M., Fedosenko Yu.</i>	
About potential of cluster schemas for optimal schedule synthesis for models of water transport logistics	421
<i>Saratov A.</i>	
Greedy algorithm for solving the NP-hard problem of minimizing total tardiness for a single machine	426
Data mining in information security	429
<i>Morozov A., Sushkova O.</i>	
Development of a Publicly Available Terahertz Video Dataset and a Software Platform for Experimenting with the Intelligent Terahertz Visual Surveillance	430
<i>Chigrinsky V., Matveev I.</i>	
Optimization of the multi-target multi-camera tracking system	432
<i>Grigorev A., Gneushev A.</i>	
Multiple object tracking on video via re-identification with descriptor pre-filtering based on quality assessment	434
<i>Sychugov A., Anchishkin A.</i>	
Method of immune detectors placement based on network node security risk assessment	439
Contents	443
Author index	466

Научное издание

ИНТЕЛЛЕКТУАЛИЗАЦИЯ
ОБРАБОТКИ ИНФОРМАЦИИ

Тезисы докладов
13-й Международной конференции

Подписано в печать 17.12.2020
Формат 60×84 1/8
Усл.-печ. л. 20,1. Уч.-изд. л. 21,17
Тираж 50 экз

Издатель — Российская Академия Наук

Печать — УНИД РАН

Отпечатано в экспериментальной цифровой типографии РАН

Издается по распоряжению президиума РАН
и распространяется бесплатно