

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

**Научный журнал Российской академии наук
(издается под руководством Отделения нанотехнологий
и информационных технологий РАН)**

Издается с 1989 года

Журнал выходит ежеквартально

Учредитель:

**Федеральный исследовательский центр
«Информатика и управление» Российской академии наук**

РЕДАКЦИОННЫЙ СОВЕТ

академик РАН И. А. Соколов — председатель Редакционного совета
академик РАН Г. И. Савин

академик РАН А. Л. Стемповский

член-корреспондент РАН Ю. Б. Зубарев

профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)

профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)

профессор В. Ротарь (V. Rotar, San-Diego, USA)

профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик РАН И. А. Соколов — главный редактор

профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора

д.т.н. В. Н. Захаров д.ф.-м.н. В. И. Синицын

проф., д.ф.-м.н. А. И. Зейфман проф., д.т.н. И. Н. Синицын

проф., д.т.н. В. Д. Ильин проф., д.ф.-м.н. В. Г. Ушаков

проф., д.т.н. К. К. Колин к.ф.-м.н. А. К. Горшенин — отв. секретарь

проф., д.ф.-м.н. В. Ю. Королев к.ф.-м.н. С. А. Христочевский

к.ф.-м.н. Р. В. Разумчик

Редакция

к.ф.-м.н. Е. Н. Арутюнов

к.ф.-м.н. Р. В. Разумчик

С. Н. Стригина

© Федеральный исследовательский центр «Информатика
и управление» Российской академии наук, 2020

Журнал включен в базу данных Russian Science Citation Index (RSCI),
интегрированную с Web of Science

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)

Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),
в базу данных Ulrich's periodicals directory

и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНТИ
и в системе Google Scholar

Журнал включен в сформированный Минобрнауки России Перечень рецензируемых научных
изданий, в которых должны быть опубликованы основные научные результаты диссертаций
на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

<http://www.ipiran.ru/journal/collected>

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 30 № 4 Год 2020

СОДЕРЖАНИЕ

Программное обеспечение исследований в области
статистического анализа данных

М. П. Кривенко 4

Усредненная вероятность ошибки вычисления
коэффициентов вейвлет–вейглет-разложения
при обращении преобразования Радона

А. А. Кудрявцев, О. В. Шестаков 14

Применение кластеризации в задачах размещения
подвижных точек доступа в воздушно- наземных
беспроводных сетях

**Е. Г. Медведева, Э. М. Хайров, Н. А. Поляков,
Ю. В. Гайдамака** 25

Нечеткое управление гетерогенным мышлением агентов
гибридной интеллектуальной многоагентной системы

С. В. Листопад, С. Б. Румовская 38

Статистический анализ биосигналов при изучении
зрительной рабочей памяти человека

А. В. Ерофеева, Т. В. Захарова 50

Построение моделей процесса с помощью простых сетей
Петри

**И. Ю. Терёхина, А. А. Грушо, Е. Е. Тимонина,
С. Я. Шоргин** 61

Комплексное обеспечение информационной безопасности
частных облачных вычислительных сред

**А. А. Грушо, А. В. Николаев, В. О. Писковский,
В. В. Сенчило, Е. Е. Тимонина** 76

Об алгоритмах декодирования обобщенных кодов
Рида–Соломона

С. М. Рацеев, О. И. Череватенко 83

СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 30 № 4 Год 2020

СОДЕРЖАНИЕ

Развитие гибридной многоядерной рекуррентной архитектуры на ПЛИС <i>Ю. А. Степченков, Н. В. Морозов, Ю. Г. Дьяченко, Д. В. Хилько, Д. Ю. Степченков</i>	95
Метод сегментации изображений на основе квадродерева <i>Ю. А. Маняков, А. И. Сорокин</i>	102
Программный комплекс для позиционирования акустических донных систем <i>В. А. Смирнов, Н. Н. Скворцова, Е. М. Кончеков, В. А. Ларичев, Г. А. Максимов</i>	113
Машинный перевод: индикаторная оценка результатов обучения искусственной нейронной сети <i>А. Ю. Егорова, И. М. Зацман, М. Г. Кружков, В. А. Нуриев</i>	124
Механизм темпорального сопоставления конкретно-исторических фактов <i>И. М. Адамович, О. И. Волков</i>	138
Блокчейн-технологии нормализованного бюджетного обеспечения национальных проектов <i>А. В. Ильин, В. Д. Ильин</i>	147
Повышение эффективности обработки информации нейронными сетями с использованием параболических интегрограференциальных сплайнов в качестве функций активации нейронов <i>Т. К. Бирюкова</i>	159
Об авторах	168
Авторский указатель за 2020 г.	172
2020 Author Index	179
Правила подготовки рукописей статей	185
Requirements for manuscripts	189

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ИССЛЕДОВАНИЙ В ОБЛАСТИ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ

М. П. Кривенко¹

Аннотация: При разработке статистических методов анализа данных применительно к задачам классификации наблюдений в области диагностической медицины, социологии, обработки текстов, распознавания изображений необходимо проводить исследования жизнеспособности используемых моделей данных, ставить эксперименты с данными и алгоритмами их обработки. Для этого приходится обращаться к возможностям вычислительной техники и создавать соответствующее программное обеспечение (ПО). В статье рассматривается опыт разработки варианта ПО, ориентированного на решение задач классификации данных. Приводится состав модулей, обеспечивающих построение и использование вероятностных моделей данных, оценивание их параметров, исследование эффективности предлагаемых процедур принятия решений. Также приводятся примеры решения конкретных прикладных задач и кратко описываются используемые элементы ПО. Характеризуются условия и возможные пути дальнейшего развития структуры, состава и содержания процедур статистического анализа.

Ключевые слова: программное обеспечение; статистический анализ данных; язык программирования Делфи; применение в различных областях

DOI: 10.14357/08696527200401

1 Введение

Создание статистических методов анализа данных подразумевает реализацию следующих этапов: формализация постановки задачи анализа, выбор модели данных, формирование методов обработки данных, интерпретация результатов. Для этого необходимо привлекать уже известные решения и развивать новые подходы, проводить множество экспериментов по обработке реальных и смоделированных данных, уметь анализировать и представлять в доходчивом виде результаты этих экспериментов. Комплексное решение перечисленных проблем немыслимо без применения средств вычислительной техники и соответствующего ПО обработки информации.

В тексте будут встречаться общезвестные названия пакетов и систем программ, сред программирования и т. п., по которым в интернете можно легко найти сведения об их назначении и составе.

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, mkrivenko@ipiran.ru

Исследовательский характер построения моделей и методов статистического анализа данных сформировал специфические условия программирования:

- цель исследований — решение прикладных задач, а не создание ПО, поэтому мешающим фактором выступают постоянные изменения в области программирования, которые мало что дают по содержанию, но отнимают уйму времени при освоении;
- при разработке программ отдается предпочтение надежности получаемого продукта и не всегда обращается внимание на изящность решений;
- предпочтение отдается классике, при авторском программировании использовался Borland Delphi Enterprise (version 7) на Windows XP;
- важно пояснить, почему приходилось прибегать к среде программирования, а не обращаться к статистическим пакетам (например, SAS) или близким к ним продуктам (например, Matlab); опыт показывает, что подобные пакеты становятся крайне громоздкими, когда речь заходит о поиске новых моделей и методов, не всегда предоставляют четкие обоснования для применяемых решений; не надо забывать и о финансовых проблемах; понятно, что последнее не означает, что не нужно использовать такое мощное подспорье, как разработанные математические пакеты (например, LAPACK);
- постоянное изменение среды программирования (в частности, операционные системы вплоть до Window 10), что в итоге привело к необходимости осваивать новые средства (кстати, попытки пойти путем создания на Window 10 вариантов виртуальных машин под Windows XP для Delphi 7 ни к чему полезному не привели), а также требуют модернизации разработанного ПО.

2 Структура и общая характеристика программного обеспечения

Программное обеспечение состоит из следующих групп модулей:

- процедуры обработки данных;
- динамически подключаемая библиотека процедур на основе пакета LAPACK;
- программы, обеспечивающие решение отдельных задач анализа данных.

При описании модуля обработки данных будем следовать схеме: факторы создания, общая характеристика состава. При изложении будем различать тривиальные решения при создании программных объектов, решения на основе известных идей (частично оригинальные) и оригинальные решения.

Модуль Common: общие определения и процедуры, а также формы интерфейса. Факторы создания: базовые решения, фиксация приемов программирования. В состав модуля входят, в основном, тривиальные объекты, например: константа — значение порога длины сортируемой последовательности для выбора наиболее эффективного алгоритма сортировки; процедура упорядочения массива чисел с удалением повторяющихся чисел или процедура нахождения

рангов; класс для представления «числа» в смешанной системе счисления. Но присутствуют и частично оригинальные решения, например вычисление логарифма гамма-функции (на его основе удалось обеспечить с заданной точностью вычисление комбинаторных формул). Отдельную группу составляют модули, включающие формы для ввода-вывода информации, например работа с масштабируемым изображением.

Модуль MatrixComp: матричные операции и процедуры линейной алгебры. Факторы создания: обеспечить возможность работы с базовыми элементами линейной алгебры. Понятно, что по составу часть объектов являются тривиальными, например простейшие преобразования матриц. При этом число различных действий с матрицами достаточно велико, например для класса обычных вещественных матриц определено около 60 процедур, охватывающих изменение структуры матрицы до ее сингулярного разложения. В этом модуле появилась необходимость обращения к пакету LAPACK; примером может служить функция dsyev.c на языке C, подключаемая как DLL.

Модуль RandomGen: генерация псевдослучайных объектов. Факторы создания: обеспечить обработку и моделирование данных базовыми датчиками. В основном, модуль содержит тривиальные объекты. Примером оригинальной процедуры может служить генератор дискретного распределения на основе смеси двухточечных распределений.

Модуль ProbDistrib: вычисления значений функций распределения и плотности распределения. Факторы создания: обеспечить реализацию критериев значимости. Модуль преимущественно содержит объекты, основанные на известных соотношениях для специальных функций. Много кропотливого труда пришлось потратить на построение методов вычислений для распределений, использующих комбинаторные элементы, например в случае биномиального и гипергеометрического распределений, основанных на числах Стирлинга; это привело к созданию полностью оригинального продукта.

Модуль Statistics: нахождение различных статистик. Факторы создания: обеспечить обработку данных базовыми объектами. Модуль содержит тривиальные объекты, например преобразование данных с помощью перехода к главным компонентам или к проекциям на левые сингулярные векторы матрицы наблюдений, варианты процедур оценивания выборочных моментов для многомерных данных, процедуру сравнения вероятностей успеха в двух последовательностях испытаний Бернулли. Примерами частично оригинальных объектов могут служить перечисление всех разбиений множества элементов (реализован алгоритм 1.19 [1]) или критерий Сю (Хи) [2] однородности биномиальных пропорций. Интересна процедура сравнения вероятностей успеха в двух последовательностях испытаний Бернулли как пример демонстрации оригинальных приемов работы с комбинаторными формулами и как пример руководства для практического применения.

Модуль NormMixt: работа со смесью нормальных распределений. Факторы создания: обеспечить возможность работы с базовой моделью данных.

Практически все объекты, входящие в состав этого модуля, — оригинальные. Рассматриваются случаи одномерного и многомерных распределений, представлены вариации алгоритмов, нацеленные на улучшение временных и емкостных характеристик программ. Ярко проявляется тенденция к появлению тупиковых вариантов программ, которые «раздувают» проект, а выбросить «жалко».

Модуль ImageExec: работа с изображениями. Факторы создания: поддержка решения задач распознавания элементов изображения. Все разработанные объекты — оригинальные. Многие из них помечены как экспериментальные, реализуют разведочные (поисковые) шаги в анализе изображений.

Модуль GenPCA: анализ главных компонент. Факторы создания: обеспечить расширение принятой модели данных и снижение размерности. Все объекты — оригинальные.

Модуль Estimators: нахождения различных статистик. Факторы создания: дополнить перечень необходимых статистик. Большинство объектов являются частично оригинальными.

3 Применение программного обеспечения на практике

Рассмотрим примеры реальных задач анализа данных, определивших состав и содержание ПО в последнее десятилетие. Соответствующие материалы авторские и опубликованы в печати.

Задача выборочного контроля. Предложенные решения позволили обеспечить реализацию процесса выборочного контроля объектов как в случае полной исходной информации, так и частично имеющейся. Учет точности вычислений привел к проработке процедур расчета комбинаторных формул.

Задача построения эмпирического байесовского классификатора объектов разной размерности. Представлен комбинированный метод построения оценок элементов байесовского классификатора, включающих непараметрическую ядерную оценку и параметрическую оценку плотности распределения данных. Разработаны и исследованы соответствующие процедуры построения оценок, выбора ядра и его параметров. Продуктивность рассмотренных идей иллюстрируется на примере распознавания реальных сильно зашумленных изображений текстов в случае, когда отдельные знаки имеют разные размеры. Предложенная комбинированная оценка плотностей распределений и ее экспериментальный анализ показали высокую эффективность байесовского подхода к классификации объектов, имеющих разную размерность, а также работоспособность в случае малых объемов обучающей выборки.

Задача предварительной обработки изображений текста, включающей коррекцию наклона и выделения строк; при этом распознаваемое изображение обладает низким качеством и получено с высоким разрешением (обработка данных большой размерности). Приходится отказаться от бинаризации исходного изображения, усложнять и вводить новые приемы обработки данных: использование модели смеси распределений при анализе проекционных профилей,

введение поэтапного оценивания неизвестных характеристик строчной структуры текста, фрагментацию изображения, анализ локальных взаимосвязей яркости пикселов, что привело к созданию оригинального ПО.

Задача разработки методов исследования однородности статистических данных о химическом составе камней при уролитиазе. Многомерность данных, бедный спектр различных значений показателей, невозможность использовать простые адекватные вероятностные модели для них, а также отсутствие накопленного опыта подобного моделирования в области урологии придали специфику рассматриваемой задаче. При анализе изменения состава камней во времени предложен критерий, использующий парные сравнения разных фрагментов данных. В этом случае впервые применялись идеи аппроксимации возникающих связей между этими фрагментами с помощью отношения эквивалентности, что позволило обоснованно проводить стратификацию исходных данных. В ходе экспериментов с помощью оригинального ПО получены статистически обоснованные результаты как об особенностях мочекаменной болезни, так и об изменениях этих особенностей с течением времени.

Задача сравнительного анализа процедур регрессионного анализа. В перечень сравниваемых процедур были включены как обычные методы оценивания параметров модели, так и методы на основе знаковых и ранговых статистик. Реализация метода наименьших квадратов на основе знаковых статистик потребовала построения эффективной процедуры обработки данных, для чего был описан и исследован вариант метода наискорейшего спуска. В сложной ситуации кусочно-постоянной целевой функции получен работоспособный вариант соответствующего алгоритма. Проведенный сравнительный анализ процедур регрессионного анализа в реальных условиях, когда данные не подчиняются нормальному распределению, показал преимущества непараметрических методов. Достоинства непараметрических методов позволили повысить точность согласования двух способов измерения хромогранина А, широко используемого в качестве иммуногистохимического маркера нейроэндокринной дифференцировки.

Задача отбора признаков классификации. Для оценивания качества классификации использовались метод перепроверки и бутстреп-метод, а для исследования ряда выборочных значений — сравнительный анализ доверительных интервалов и критерии однородности биномиальных пропорций. Это потребовало создания ПО поддержки процессов повторной обработки частей выборки, а также приемов точного вычисления комбинаторных формул. Проведенные эксперименты в задаче прогнозирования химического состава мочевых камней создают предпосылки для повышения качества получаемых решений (сокращение перечня проводимых анализов может привести к росту точности прогноза), дают толчок специалистам в предметной области для прояснения сути протекающих процессов.

Задача построения слабых критериев значимости на основе многомерных референсных регионов (MRR, multivariate reference regions) примени-

тельно к области медицинской диагностики. Предложен оригинальный метод построения региона на основе областей с высокой плотностью точек и аппроксимации распределения данных с помощью смеси нормальных распределений. Для оценки порога для плотности распределения использовался бутстреп-метод. В качестве эксперимента рассматривалось построение и использование эталонной области для прогнозирования типа мочевого камня. Обработка реальных данных продемонстрировала преимущества предлагаемых решений. На настоящий момент имеется относительно мало примеров применения MRR в клинической практике. Тому есть несколько причин, и одна из них состоит в том, что лаборатории слабо оснащены программными средствами для реализации достаточно сложного математического аппарата многомерного анализа, а еще важнее, что отсутствуют методики, инструкции по использованию соответствующих средств. Созданное ПО работы с MRR может послужить прототипом для последующих разработок. Уже проведенные на его базе эксперименты позволяют сформулировать некоторые рекомендации по исследованию и разработке MRR.

Задача обучаемой классификации неполных данных. Обучение байесовского классификатора в этом случае проводилось с помощью метода максимального правдоподобия для модели смеси нормальных распределений. Строгий вывод формул, обеспечивающих реализацию шагов EM (expectation-maximization) алгоритма, позволил корректно применять итерационный процесс получения оценок параметров смеси. Для неполных данных предложены приемы выбора начальных значений и коррекции вырождающихся ковариационных матриц элементов смеси. Экспериментальная часть работы заключалась в анализе зависимости качества классификации от степени пропуска отдельных значений, для этого использовались реальные данные о ферментах, полученные для пациентов с заболеваниями печени. Их обработка продемонстрировала практически идентичные ошибки классификации при применении простых и сложных методов обработки пропусков в случае невысокой степени случайного пропуска отдельных значений.

Задача сжатия данных. Здесь новые возможности открывает вероятностный анализ главных компонент (PPCA, probabilistic principal component analysis), реализуемый в рамках принципа максимального правдоподобия для гауссовской модели с латентными переменными. В рамках PPCA появились алгоритмы обработки данных, нацеленные на снижение размерности данных и обеспечивающие переход в пространство главных компонент, но не дающие в явном виде характеристики главных компонент. В ходе авторских исследований известные результаты были дополнены деталями, которые углубляли понимание особенностей PPCA, имели прикладное значение, исправляли выявленные ошибки в публикациях. Предложены и обоснованы два метода реконструкции характеристик главных компонент. Они позволяют воспользоваться результатами предельного PPCA и проводить полный анализ данных, который особенно эффективен при существенном снижении размерности.

Задача обучаемой классификации данных с учетом анализа главных компонент (PCA, principal component analysis). Построение байесовского классификатора становится возможным после представления ковариаций через параметры вероятностной модели РСА. Проведено исследование качества классификации с учетом реальной размерности данных. Продемонстрировано, что при ее правильном задании классификатор обладает наименьшими вероятностями ошибки. При этом превышение наилучшего значения размерности обычно ухудшает качество классификации в меньшей степени, чем его занижение. Смесь вероятностных анализаторов главных компонент позволяет моделировать объемные данные с помощью относительно небольшого числа свободных параметров. Число свободных параметров можно контролировать с помощью выбора латентной размерности данных. Получить перечисленные содержательные выводы удается методом моделирования, что подразумевает создание специфического ПО, насыщенного множеством деталей (например, приходится выделять случай сингулярных распределений данных, для которого оценивание параметров модели проводится при ограничениях на собственные значения ковариационных матриц).

Задача выбора модели данных. Эффективное решение задач медицинской диагностики требует применения сложных вероятностных моделей, позволяющих адекватно описывать реальные данные и допускающих применение аналитических методов обучаемой классификации. Выбор модели смеси нормальных распределений решает поставленные задачи, но приводит к проблеме проклятия размерности. Переход к модели смеси вероятностных анализаторов главных компонент позволяет формально поставить задачу выбора ее структурных параметров. Решение предлагается искать, комбинируя применение информационных критерий для формирования начальных приближений с последующим уточнением получающихся оценок. На примере экспериментов по диагностированию болезней печени и прогнозированию химического состава мочевых камней демонстрируются возможности описанных процедур анализа данных. Предлагаемые решения требуют создания оригинального ПО, но при этом становятся источником повышения точности классификации, дают толчок специалистам в предметной области для прояснения сути протекающих процессов.

Задача выбора структурных параметров, характеризующих модель смеси вероятностных анализаторов главных компонент, а именно: числа элементов смеси и размерностей этих элементов. Среди набора используемых на практике подходов в случае задачи обучаемой классификации данных фактически доступными остаются лишь методы управления выборкой. Для реализации выбора размерностей приходится использовать комбинацию известных методов выбора размерностей принятой модели. Но, чтобы получить итоговый результат о параметрах модели, приходится строить программную реализацию всех вариантов и проводить емкие по времени исследования, для чего также нужно создавать ПО и только после этого делать вывод о наилучшей модели для решаемой задачи. На примере экспериментов по диагностированию болезней печени и прогно-

зированию химического состава мочевых камней демонстрируются достоинства описанных процедур анализа данных.

Задача анализа результатов наблюдения группы объектов в течение времени, за которое эти объекты могут поменять какие-либо свои значимые характеристики. Цель состоит в том, чтобы описать изменения и выявить факторы, их определяющие. Соответствующие методы известны как лонгитюдные (продольные, протяженные во времени). В ходе исследований предложен иной подход — консолидирующий, когда серия многомерных характеристик некоторого объекта составляет единый вектор наблюденных значений. За счет увеличения размерности данных удается получить единую картину описания объектов, формализовать построение модели данных. Для демонстрации сути подхода и иллюстрации появляющихся возможностей анализа данных рассматривалась задача ранней диагностики рака с использованием биомаркера PSA. Выявлено, что многомерный подход при анализе серий анализов приводит к повышению точности диагностики. При применении консолидирующего подхода, безусловно, возрастает аналитическая и вычислительная трудоемкость методов решения возникающих задач (оценивание параметров модели данных, снижение размерности пространства признаков, реализация решающих процедур), что усиливает роль создаваемого ПО.

4 Заключение

Перевод разработанных программ на современный язык Embarcadero Delphi 10.3 выявил ряд моментов, требующих дополнительного внимания.

Выбор и «внедрение» в текст программ обновленных типов данных. Если для числовых данных речь идет о контроле точности представления и времени обработки данных и достаточно проверить интегральные показатели программ решения отдельных комплексных задач, то в случае строковых данных приходится верифицировать все фрагменты работы со строками, чтобы исключить случаи неправильной работы с этими объектами.

Использование возможностей задания платформы (32/64 bit Intel Windows), что влияет на точность вычислений и возможности обработки данных большой размерности. Основным фактором здесь оказывается проблема проклятия размерности в том смысле, что переход к 64-битной платформе может привести к нереализуемой возможности работы с многомерными данными большого объема (в первую очередь речь идет о применении методов управления обработкой выборкой). К сожалению, единственный путь прояснить это состоит в реализации обработки данных на обеих платформах и в проведении экспериментов с решением конкретных задач.

Освоение средств распараллеливания алгоритмов, которые заявлены авторами программной среды Delphi и требуют апробации. Не надо забывать следующее. При активном использовании моделирования в процессе исследований свойств статистических процедур приходится обращаться к генераторам

случайных объектов, что приводит к необходимости одновременного использования ряда программных датчиков с определенными требованиями к ним. А это, в свою очередь, требует восстановить имеющиеся алгоритмы, разработать и запрограммировать новые.

Уточнение состава модулей, так как постоянное наполнение ПО новыми возможностями приводит подчас к повторному программированию и затруднениям при поиске необходимых средств анализа данных. Кроме того, необходимость исключить ошибки программирования может приводить к искусственному усложнению структуры модулей (примером может служить появление двух функционально подобных модулей — Statistics и Estimators).

Систематизация использования сторонних математических библиотек, что в первую очередь касается LAPACK.

Уточнение источников математических и алгоритмических решений, использованных при создании ПО.

Литература

1. Липский В. Комбинаторика для программистов. — М.: Мир, 2011. 213 с.
2. Klein M., Linton P. On a comparison of tests of homogeneity of binomial proportions // J. Statistical Theory Applications, 2013. Vol. 12. Iss. 3. P. 208–224.

Поступила в редакцию 31.08.20

SOFTWARE OF RESEARCH IN STATISTICAL DATA ANALYSIS

M. P. Krivenko

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: When developing statistical methods for data analysis in relation to the problems of observation classification in the field of diagnostic medicine, sociology, word processing, and image recognition, it is necessary to conduct studies of the viability of the data models used, to experiment with data and algorithms for their processing. To do this, one has to turn to the capabilities of computer technology and create the appropriate software. The article discusses the experience of developing a version of software aimed at solving data classification problems. The composition of the modules is given, which ensures the construction and use of probabilistic data models, the estimation of their parameters, and the study of the effectiveness of the proposed decision-making procedures. It also provides examples of solving specific applied problems and briefly describes the software elements used. The conditions and possible ways of further development of the structure, composition, and content of statistical analysis procedures are characterized.

Keywords: software; statistical data analysis; Delphi; applications in various areas

DOI: 10.14357/08696527200401

References

1. Lipskiy, V. 2011. *Kombinatorika dlya programmistov* [Combinatorics for programmers]. Moscow: Mir. 213 p.
2. Klein, M., and P. Linton. 2013. On a comparison of tests of homogeneity of binomial proportions. *J. Statistical Theory Applications* 12(3):208–224.

Received August 31, 2020

Contributor

Krivenko Michail P. (b. 1946) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mkrivenko@ipiran.ru

УСРЕДНЕННАЯ ВЕРОЯТНОСТЬ ОШИБКИ ВЫЧИСЛЕНИЯ КОЭФФИЦИЕНТОВ ВЕЙВЛЕТ–ВЕЙГЛЕТ-РАЗЛОЖЕНИЯ ПРИ ОБРАЩЕНИИ ПРЕОБРАЗОВАНИЯ РАДОНА*

A. A. Кудрявцев¹, O. V. Шестаков²

Аннотация: Методы реконструкции изображений, основанные на разложении функции изображения по специальному вейвлет-базису и последующей пороговой обработке коэффициентов разложения, используются при решении задач вычислительной томографии. Их привлекательность заключается в адаптации к пространственным неоднородностям изображений и возможности реконструкции локальных участков изображения по неполным проекционным данным, что имеет ключевое значение, например, для медицинских приложений, где пациента нежелательно подвергать лишней дозе облучения. Анализ погрешностей этих методов представляет собой важную практическую задачу, поскольку позволяет оценить качество как самих методов, так и используемого оборудования. В работе рассматривается метод вейвлет–вейглет-разложения при реконструкции томографических изображений в модели с аддитивным гауссовым шумом. Оценивается порядок функции потерь, основанный на усредненной вероятности ошибки вычисления вейвлет-коэффициентов.

Ключевые слова: преобразование Радона; вейвлет–вейглет-разложение; пороговая обработка; функция потерь

DOI: 10.14357/08696527200402

1 Введение

Томографические методы реконструкции изображений широко распространены в самых разных областях, включая медицину, геофизику, астрономию и радиолокацию. Математическая модель, лежащая в основе этих методов, как правило, основана на так называемом преобразовании Радона. Задача обращения этого преобразования относится к некорректно поставленным, и для ее решения

*Работа выполнена при финансовой поддержке РФФИ (проект № 19-07-00352); исследования проводились в рамках программы Московского центра фундаментальной и прикладной математики.

¹Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, nubigena@mail.ru

²Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, oshestakov@cs.msu.su

требуется использовать методы регуляризации, поскольку в реальных наблюдениях всегда присутствует шум. Классические линейные методы регуляризации не всегда дают удовлетворительные результаты, поскольку не учитывают пространственную неоднородность изображений. С этим недостатком позволяют справиться методы вейвлет-разложения, предложенные в работе [1]. В работах [2–4] проводится анализ среднеквадратичного риска методов нелинейной регуляризации, основанных на пороговой обработке коэффициентов вейвлет-разложения. В данной работе рассматривается функция потерь, основанная на вероятностях ошибок вычисления коэффициентов вейвлет-разложения, оценивается ее порядок и вычисляются асимптотически оптимальные значения порогов.

2 Обращение преобразования Радона с помощью метода вейвлет–вейглет-разложения

Преобразование Радона на плоскости определяется как набор интегралов от функции f по всевозможным прямым:

$$Rf(s, \theta) = \int_{L_{s,\theta}} f(x, y) dl,$$

где

$$L_{s,\theta} = \{(x, y) : x \cos \theta + y \sin \theta - s = 0\}.$$

Задача реконструкции томографического изображения состоит в обращении преобразования Радона. Для решения этой задачи воспользуемся методом вейвлет–вейглет-разложения, предложенным в работе [1], который позволяет учитывать различные локальные особенности изображений.

Пусть задана масштабирующая функция $\phi(x)$ и вейвлет-функция $\psi(x)$. Определим функции:

$$\left. \begin{array}{l} \psi_{j,k_1,k_2}^{[1]}(x, y) = 2^j \phi(2^j x - k_1) \psi(2^j y - k_2); \\ \psi_{j,k_1,k_2}^{[2]}(x, y) = 2^j \psi(2^j x - k_1) \phi(2^j y - k_2); \\ \psi_{j,k_1,k_2}^{[3]}(x, y) = 2^j \psi(2^j x - k_1) \psi(2^j y - k_2), \end{array} \right\} \quad (1)$$

где $j, k_1, k_2 \in \mathbb{Z}$. Семейство $\{\psi_{j,k_1,k_2}^{[\lambda]}\}$ образует ортонормированный базис в $L^2(\mathbb{R}^2)$. Индекс j в (1) называется масштабом, а индексы k_1 и k_2 — сдвигами. Функция ψ должна удовлетворять определенным требованиям, однако ее можно выбрать таким образом, чтобы она обладала некоторыми полезными свойствами, например была дифференцируемой нужное число раз, имела заданное число нулевых моментов и достаточно быстро убывала на бесконечности вместе со

своими производными [5]. Везде далее предполагается, что ψ удовлетворяет всем необходимым требованиям.

Вейвлет-разложение функции f имеет вид:

$$f = \sum_{\lambda, j, k_1, k_2} \left\langle f, \psi_{j, k_1, k_2}^{[\lambda]} \right\rangle \psi_{j, k_1, k_2}^{[\lambda]}. \quad (2)$$

Метод реконструкции f заключается в замене скалярных произведений в (2) на величины, выражающиеся через Rf , а не через f . Пусть

$$\xi_{j, k_1, k_2}^{[\lambda]} = \frac{1}{4\pi} I^{-1} R \psi_{j, k_1, k_2}^{[\lambda]},$$

где I^α — потенциал Рисса, определяемый через преобразование Фурье по формуле:

$$\widehat{I^\alpha g}(\omega) = |w|^{-\alpha} \widehat{g}(\omega).$$

Если ϕ и ψ достаточно гладкие и достаточно быстро убывают на бесконечности, то последовательность функций $\{2^{-j/2} \xi_{j, k_1, k_2}^{[\lambda]}\}$ образует устойчивый базис [2]. При этом выполнено:

$$\left\langle f, \psi_{j, k_1, k_2}^{[\lambda]} \right\rangle = \left\langle Rf, \xi_{j, k_1, k_2}^{[\lambda]} \right\rangle.$$

Функции $2^{-j/2} \xi_{j, k_1, k_2}^{[\lambda]}$ получили название «вейвлеты». Формула обращения в методе вейвлет–вейвлет-разложения выглядит следующим образом:

$$f = \sum_{\lambda, j, k_1, k_2} \left\langle Rf, \xi_{j, k_1, k_2}^{[\lambda]} \right\rangle \psi_{j, k_1, k_2}^{[\lambda]}.$$

3 Модель проекционных данных и функция потерь

На практике функция f задана в дискретных отсчетах на некотором конечном носителе. Без ограничения общности будем считать, что этим носителем выступает круг единичного радиуса с центром в начале координат. Проекционные данные в этом случае измеряются при $(s, \theta) \in [-1, 1] \times [0, \pi]$. Пусть $\{e_{i,j}, i, j \in \mathbb{Z}\}$ имеют нормальное распределение с нулевым средним и дисперсией σ^2 . Рассмотрим следующую модель проекционных данных с шумом:

$$Y_{i,j} = Rf_{i,j} + e_{i,j}, \quad i = 1, \dots, 2^J, \quad j = 1, \dots, 2^J,$$

где 2^J — число отсчетов.

При переходе к дискретному вейвлет–вейглет-преобразованию получается следующая модель дискретных коэффициентов [6]:

$$X_{j,k_1,k_2}^{[\lambda]} = \mu_{j,k_1,k_2}^{[\lambda]} + e_{j,k_1,k_2}^{[\lambda]}.$$

Здесь

$$\mu_{j,k_1,k_2}^{[\lambda]} = 2^J \left\langle Rf, \xi_{j,k_1,k_2}^{[\lambda]} \right\rangle,$$

а $e_{j,k_1,k_2}^{[\lambda]}$ уже не являются независимыми и имеют нормальное распределение с нулевым средним и дисперсией $\sigma_{\lambda,j}^2 = C_\lambda \cdot 2^j$, где константы C_λ ($\lambda = 1, 2, 3$) зависят от выбранного вейвлет-базиса.

Для построения оценок коэффициентов $\mu_{j,k_1,k_2}^{[\lambda]}$ обычно применяется пороговая обработка эмпирических коэффициентов $X_{j,k_1,k_2}^{[\lambda]}$, идея которой заключается в удалении достаточно маленьких коэффициентов, которые считаются шумом. Такой подход оправдывается свойством вейвлет-разложения экономно представлять искомую функцию, т. е. «полезных» коэффициентов немного и они достаточно велики по абсолютному значению. Характер пороговой обработки определяется пороговой функцией $\rho_{T_{\lambda,j}}(x)$ и порогом $T_{\lambda,j}$. Вместо аргумента x в функцию $\rho_{T_{\lambda,j}}(x)$ подставляются эмпирические коэффициенты $X_{j,k_1,k_2}^{[\lambda]}$, и получаются оценки коэффициентов

$$\hat{X}_{j,k_1,k_2}^{[\lambda]} = \rho_{T_{\lambda,j}} \left(X_{j,k_1,k_2}^{[\lambda]} \right).$$

Самыми распространенными являются функции жесткой пороговой обработки

$$\rho_{T_{\lambda,j}}^{(h)}(x) = x \mathbf{1}(|x| > T_{\lambda,j})$$

и мягкой пороговой обработки

$$\rho_{T_{\lambda,j}}^{(s)}(x) = \text{sign}(x)(|x| - T_{\lambda,j})_+.$$

В качестве функции потерь в методах пороговой обработки обычно рассматривается среднеквадратичный риск [5]. В работах [2–4] проводится анализ среднеквадратичного риска и его оценки, а также вычисляются оптимальные значения порогов, ориентированные на различные классы функций. В данной работе рассматривается функция потерь, основанная на вероятностях того, что ошибка вычисления коэффициента разложения превысит заданный критический уровень. Такая функция потерь была впервые предложена в работе [7], и было показано, что получаемые в этом случае оценки функции сигнала оказываются

сравнимыми, а иногда и лучшими, чем оценки, ориентированные на минимизацию среднеквадратичного риска. В работах [8, 9] оцениваются порядки этой функции в задачах прямого наблюдения сигнала и задачах обращения линейных однородных операторов.

Пусть (ξ, η, ζ) — трехмерная случайная величина, не зависящая от всех $e_{j,k_1,k_2}^{[\lambda]}$ и имеющая дискретное равномерное распределение на множестве индексов $j = 0, \dots, J - 1, k_1, k_2 = 0, \dots, 2^j - 1$. Для заданного критического уровня $\varepsilon > 0$ определим функцию потерь

$$r_J(f) = \mathbb{E}\mathbb{P}\left(\left|\hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]}\right| > \varepsilon \mid \xi, \eta, \zeta\right) = \\ = \frac{\sum_{j=0}^{J-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathbb{P}\left(\left|\hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]}\right| > \varepsilon\right)}{2^{2J}},$$

т. е. $r_J(f)$ представляет собой усредненную вероятность того, что ошибка вычисления вейвлет-коэффициента превысит критический уровень ε .

Рассмотрим $\text{Lip}(\gamma, L)$ — класс кусочно-регулярных по Липшицу функций, где $\gamma > 0$ — показатель, а $L > 0$ — константа Липшица [5]. Асимптотически оптимальный порог минимизирует порядок функции потерь

$$R_J = \sup_{f \in \text{Lip}(\gamma, L)} r_J(f), \quad (3)$$

т. е. является оптимальным в минимаксном смысле.

В следующем разделе приведены верхние и нижние оценки для асимптотически оптимального порога и функции потерь при жесткой и мягкой пороговой обработке.

4 Асимптотический порядок функции потерь

Пусть функция $g_1(J) > 0$ сколь угодно медленно убывает по J к нулю, а $g_2(J) > 0$ неограниченно возрастает по J , причем

$$\ln g_2(J) = o\left(\sqrt{\ln 2^J}\right), \quad J \rightarrow \infty.$$

Теорема. При выборе асимптотически оптимального порога для жесткой и мягкой пороговой обработки функция потерь (3) удовлетворяет неравенствам

$$C_1 \cdot 2^{-(2\gamma/(\gamma+1))J} g_1(J) \leq R_J \leq C_2 \cdot 2^{-(2\gamma/(\gamma+1))J} g_2(J),$$

где C_1 и C_2 — некоторые положительные константы. Для асимптотически оптимального значения порога, минимизирующего порядок функции потерь (3) при жесткой и мягкой пороговой обработке, начиная с некоторого J , справедливо неравенство

$$T_* - T_2 \leq T_{\lambda,j} \leq T_* - T_1,$$

где

$$\begin{aligned} T_* &= C_\lambda^{1/2} 2^{j/2} \sqrt{\frac{2\gamma}{\gamma+1} \ln 2^{2j}}; \\ T_i &= C_\lambda^{1/2} 2^{j/2} \sqrt{\frac{\gamma+1}{2\gamma}} \frac{\ln \left((\ln 2^{2j})^{1/2} g_i(j) \right)}{\sqrt{\ln 2^{2j}}}, \quad i = 1, 2. \end{aligned}$$

Доказательство. Докажем теорему для случая жесткой пороговой обработки. Случай мягкой пороговой обработки рассматривается аналогично.

Пусть

$$\hat{X}_{j,k_1,k_2}^{[\lambda]} = \rho_{T_{\lambda,j}}^{(h)} \left(X_{j,k_1,k_2}^{[\lambda]} \right).$$

Справедливо неравенство [5]:

$$\left| \mu_{j,k_1,k_2}^{[\lambda]} \right| \leq \frac{A_{\gamma,\lambda} 2^J}{2^{j(\gamma+1)}}, \quad (4)$$

где $A_{\gamma,\lambda}$ — некоторая константа, зависящая от γ и λ .

Неравенство (4) позволяет разбить все множество индексов $\{0, \dots, J-1\}$ на три класса в зависимости от величины $|\mu_{j,k_1,k_2}^{[\lambda]}|$. Пусть индексы j_1 и j_2 ($j_1 < j_2$) таковы, что

$$\left| \mu_{j,k_1,k_2}^{[\lambda]} \right| \leq \begin{cases} (g_1(J))^{-(\gamma+1)/2}, & j_1 \leq j \leq j_2 - 1; \\ (g_2(J))^{-(\gamma+1)/2}, & j_2 \leq j \leq J-1. \end{cases}$$

При этом в силу (4) для $i = 1, 2$

$$j_i = \frac{J}{\gamma+1} + \frac{\log_2 g_i(J)}{2} + \frac{\log_2 A_{\gamma,\lambda}}{\gamma+1}. \quad (5)$$

Разобьем сумму в $r_J(f)$ на три составляющие:

$$\begin{aligned}
 r_J(f) &= \sum_{j=0}^{j_1-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) + \\
 &\quad + \sum_{j=j_1}^{j_2-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) + \\
 &\quad + \sum_{j=j_2}^{J-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) \equiv \\
 &\equiv S_1 + S_2 + S_3. \tag{6}
 \end{aligned}$$

Рассмотрим S_3 . Заметим, что для любого $\varepsilon > 0$ найдется такое $J_0 = J_0(\varepsilon)$, что $(g_2(J))^{-(\gamma+1)/2} \leq \varepsilon$ и $\varepsilon \leq cT_{\lambda,j}$, $0 < c < 1$, $j_2 \leq j \leq J-1$, для всех $J > J_0$, и найдется такое $J_1 = J_1(\varepsilon, c) \geq J_0$, что для всех $J > J_1$ имеют место соотношения:

$$\mu_{j,k_1,k_2}^{[\lambda]} + \varepsilon \geq 0; \quad \mu_{j,k_1,k_2}^{[\lambda]} - \varepsilon \leq 0; \quad \mu_{j,k_1,k_2}^{[\lambda]} + \varepsilon \leq T_{\lambda,j}; \quad \mu_{j,k_1,k_2}^{[\lambda]} - \varepsilon \geq -T_{\lambda,j}$$

для $j_2 \leq j \leq J-1$. Следовательно, повторяя рассуждения из работы [9], для одного слагаемого из S_3 получаем при $J > J_1$

$$\mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) \asymp \frac{\sigma_{\lambda,j} \exp \left\{ -T_{\lambda,j}^2 / (2\sigma_{\lambda,j}^2) \right\}}{T_{\lambda,j}}.$$

Следовательно, существуют такие константы $c_1 > 0$ и $c_2 > 0$, что при $T_* - T_2 \leq T_{\lambda,j} \leq T_* - T_1$

$$c_1 \cdot 2^{2J/(\gamma+1)} g_1(J) \leq S_3 \leq c_2 \cdot 2^{2J/(\gamma+1)} g_2(J). \tag{7}$$

Найдем верхнюю оценку для функции потерь (3). Для этого предположим, что все слагаемые в суммах S_1 и S_2 из (6) отделены от нуля некоторой константой. Поэтому из (5) получаем, что

$$S_1 + S_2 = \sum_{j=0}^{j_2-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) \asymp 2^{2j_2} \asymp 2^{2J/(\gamma+1)} g_2(J). \tag{8}$$

Как видно, порядок (8) совпадает с порядком правой части (7). Следовательно, порог $T_* - T_2$ служит нижней границей (с точностью до величины порядка

$O(1/\sqrt{\ln 2^{2J}})$ для асимптотически оптимального в смысле функции потерь R_J порога.

Теперь найдем нижнюю границу для функции потерь (3). Заметим, что найдется такая функция $f \in \text{Lip}(\gamma, L)$, что в неравенстве (4) будет достигаться равенство для $0 \leq j \leq j_1 - 1$ [5]. Следовательно, существует такое $J_2 > 0$, что для всех $\varepsilon > 0$ и $J > J_2$ выполняется

$$\left| \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \text{ при } 0 \leq j \leq j_1 - 1.$$

Тогда, как и в работе [9], получаем:

$$\mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) \geq 2 - 2\Phi \left(\frac{\varepsilon}{C_\lambda^{1/2}} \right).$$

В этом случае порядок суммы S_1 в (6) равен числу слагаемых, т. е.

$$S_1 = \sum_{j=0}^{j_1-1} \sum_{k_1=0}^{2^j-1} \sum_{k_2=0}^{2^j-1} \mathsf{P} \left(\left| \hat{X}_{j,k_1,k_2}^{[\lambda]} - \mu_{j,k_1,k_2}^{[\lambda]} \right| > \varepsilon \right) \asymp 2^{2j_1} \asymp 2^{2J/(\gamma+1)} g_1(J).$$

Заметим, что сумма S_2 в данных рассуждениях не присутствует. Это означает, что истинное значение R_J имеет порядок не ниже данного, т. е. рассматриваемый порядок является нижней оценкой для истинного порядка функции потерь, а $T_* - T_1$ — верхней границей для асимптотически оптимального порога $T_{\lambda,j}$, поскольку, чтобы не удалить важные компоненты функции сигнала, следует выбирать наименьший порог, не ухудшающий порядок функции потерь.

Теорема доказана.

5 Заключение

В данной работе показано, что оптимальный порог в задаче обращения преобразования Радона при мягкой и жесткой пороговой обработке для функции потерь (3) имеет структуру, аналогичную структуре оптимального порога в задаче обращения линейного однородного оператора [9].

Литература

1. Donoho D. Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition // Appl. Comput. Harmon. A., 1995. Vol. 2. P. 101–126.
2. Lee N. Wavelet–vaguelette decompositions and homogenous equations. — West Lafayette, IN, USA: Purdue University, 1997. PhD Thesis. 103 p.
3. Abramovich F., Silverman B. W. Wavelet decomposition approaches to statistical inverse problems // Biometrika, 1998. Vol. 85. No. 1. P. 115–129.

4. Маркин А. В., Шестаков О. В. Асимптотики оценки риска при пороговой обработке вейвлет-вейглет коэффициентов в задаче томографии // Информатика и её применения, 2010. Т. 4. Вып. 2. С. 36–45.
5. Mallat S. A wavelet tour of signal processing. — New York, NY, USA: Academic Press, 1999. 857 p.
6. Захарова Т. В., Шестаков О. В. Теория вейвлетов и ее применение в обработке сигналов. — М.: МастерПринт, 2018. 180 с.
7. Sadasivan J., Mukherjee S., Seelamantula C. S. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising // 39th IEEE Conference (International) on Acoustics, Speech and Signal Processing Proceedings. — Piscataway, NJ, USA: IEEE, 2014. P. 4249–4253.
8. Кудрявцев А. А., Шестаков О. В. Асимптотическое поведение порога, минимизирующего усредненную вероятность ошибки вычисления вейвлет-коэффициентов // Докл. Акад. наук, 2016. Т. 468. № 5. С. 487–491.
9. Кудрявцев А. А., Шестаков О. В. Минимизация ошибок вычисления вейвлет-коэффициентов при решении обратных задач // Информатика и её применения, 2018. Т. 12. Вып. 2. С. 17–23.

Поступила в редакцию 18.08.20

AVERAGE PROBABILITY OF ERROR IN CALCULATING WAVELET–VAGUELETTE COEFFICIENTS WHILE INVERTING THE RADON TRANSFORM

A. A. Kudryavtsev¹ and O. V. Shestakov^{1,2}

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

²Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: Image reconstruction methods based on decomposition of the image function in a special wavelet basis and subsequent thresholding of the decomposition coefficients are used to solve computational tomography problems. Their attractiveness lies in adaptation to spatial inhomogeneities of images and the possibility of reconstructing local areas of the image from incomplete projection data that is of key importance, for example, for medical applications where it is undesirable to expose a patient to an unnecessary dose of radiation. The analysis of errors of these methods is an important practical task, since it allows one to

assess the quality of both the methods themselves and the equipment used. The paper considers the wavelet–vaguelette decomposition method for reconstructing tomographic images in a model with an additive Gaussian noise. The order of the loss function based on the average probability of error in calculating wavelet coefficients is estimated.

Keywords: Radon transform; wavelet–vaguelette decomposition; thresholding; loss function

DOI: 10.14357/08696527200402

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project No. 19-07-00352). The research was conducted in accordance with the program of the Moscow Center for Fundamental and Applied Mathematics.

References

1. Donoho, D. 1995. Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition. *Appl. Comput. Harmon. A.* 2:101–126.
2. Lee, N. 1997. Wavelet–vaguelette decompositions and homogenous equations. West Lafayette, IN: Purdue University. PhD Diss. 103 p.
3. Abramovich, F., and B. W. Silverman. 1998. Wavelet decomposition approaches to statistical inverse problems. *Biometrika* 85(1):115–129.
4. Markin, A. V., and O. V. Shestakov. 2010. Asimptotiki otsenki riska pri porogovoy obrabotke veyvlet–veyglet koeffitsientov v zadache tomografii [The asymptotics of risk estimate for wavelet–vaguelette coefficients' thresholding in the problems of tomography]. *Informatika i ee Primeneniya — Inform. Appl.* 4(2):36–45.
5. Mallat, S. 1999. *A wavelet tour of signal processing*. New York, NY: Academic Press. 857 p.
6. Zakhарова, Т. В., and О. В. Шестаков. 2016. *Teoriya veyvletov i ee primenie v obrabotke signalov* [Wavelet theory and its application in signal processing]. Moscow: MasterPrint. 180 p.
7. Sadashivan, J., S. Mukherjee, and C. S. Seelamantula. 2014. An optimum shrinkage estimator based on minimum-probability-of-error criterion and application to signal denoising. *39th IEEE Conference (International) on Acoustics, Speech and Signal Processing Proceedings*. Piscataway, NJ: IEEE. 4249–4253.
8. Kudryavtsev, A. A., and O. V. Shestakov. 2016. Asymptotic behavior of the threshold minimizing the average probability of error in calculation of wavelet coefficients. *Dokl. Math.* 93(3):295–299.
9. Kudryavtsev, A. A., and O. V. Shestakov. 2018. Minimizatsiya oshibok vychisleniya veyvlet-koeffitsientov pri reshenii obratnykh zadach [Minimization of errors of calculating wavelet coefficients while solving inverse problems]. *Informatika i ee Primeneniya — Inform. Appl.* 12(2):17–23.

Received August 18, 2020

Contributors

Kudryavtsev Alexey A. (b. 1978) — Candidate of Sciences (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; nubigena@mail.ru

Shestakov Oleg V. (b. 1976) — Doctor of Sciences in physics and mathematics, professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; oshestakov@cs.msu.su

ПРИМЕНЕНИЕ КЛАСТЕРИЗАЦИИ В ЗАДАЧАХ РАЗМЕЩЕНИЯ ПОДВИЖНЫХ ТОЧЕК ДОСТУПА В ВОЗДУШНО-НАЗЕМНЫХ БЕСПРОВОДНЫХ СЕТЯХ*

Е. Г. Медведева¹, Э. М. Хайров², Н. А. Поляков³, Ю. В. Гайдамака⁴

Аннотация: Выполнен обзор задач, возникающих в беспроводных сетях с подвижными базовыми станциями, размещенными на беспилотных летательных аппаратах (БПЛА). Сформулирована и решена задача оптимизации для определения позиций БПЛА, обеспечивающих максимизацию вероятности покрытия в области предоставления связи при ограничениях на интерференцию от соседних базовых станций. Для сравнительного анализа эффективности развертывания сети при применении адаптивной навигации выбраны два метода — метод k-средних и метод роя частиц, в основе которых лежит кластеризация пользователей. Применение методов проиллюстрировано для сценария обеспечения связью участников массового мероприятия на открытой местности (сценарий «концерт»).

Ключевые слова: БПЛА; воздушно- наземная сеть; метод роя частиц; метод k-средних; вероятность покрытия

DOI: 10.14357/08696527200403

1 Введение

Для преодоления кризиса дефицита спектра частот, связанного с экспоненциальным ростом беспроводных устройств, одним из возможных решений представляется расширение спектра до более высоких диапазонов, в которых по-прежнему доступно большое число лицензированных и нелицензированных частот. Многообещающим подходом стало повсеместное внедрение и использование частот миллиметровых волн (mmWave) в диапазоне от 30 до 300 ГГц [1]. С этим несравненным преимуществом технология mmWave стала важной компонентой

* Публикация подготовлена при поддержке Программы РУДН «5-100». Исследование выполнено при финансовой поддержке РФФИ (проекты № 18-07-00576 и № 20-07-01064).

¹Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, medvedeva-eg@rudn.ru

²Российский университет дружбы народов, emil.khayrov@gmail.com

³Российский университет дружбы народов, goto97@mail.ru

⁴Российский университет дружбы народов; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, gaydamaka-yuv@rudn.ru

в беспроводных системах 5G и последующих сетях 5G (beyond 5G, B5G), характеризующихся требованиями к высокой пропускной способности, широкой полосе пропускания, сверхбыстрым скоростям передачи данных, сверхнизкой задержке и расширенным возможностям подключения для сетевого сообщества. В будущих сетях 5G подразумевается поддержка большого разнообразия широкодиапазонных приложений, предусмотренных для множественных / многоканальных беспроводных систем (Multiple Gigabit Wireless Systems, MGWS), развернутых на нелицензированных частотах диапазона около 60 ГГц, зарезервированного для промышленных, научных и медицинских целей, и сетей WiGig (беспроводной гигабитной) как продолжение сетей Wi-Fi в диапазоне 60 ГГц. Кроме того, использование миллиметрового диапазона длин волн выделено в качестве ключевого решения сетей 5G «Новое радио» (New Radio, NR), разработанного консорциумом 3GPP в качестве унифицированного радиоинтерфейса для поддержки различных пользовательских услуг, таких как, например, усовершенствованная мобильная широкополосная связь (Enhanced Mobile Broadband, eMBB) и сверхнадежная связь с низкой задержкой (Ultra-Reliable Low Latency Communication, URLLC).

Диапазон электромагнитных волн в миллиметровом диапазоне, называемый также диапазоном сверхвысоких частот, определяется полосой спектра от 30 до 300 ГГц, между микроволнами и инфракрасными волнами. Наличие обильной незанятой полосы пропускания, доступной на частотах mmWave, — одно из ключевых преимуществ связи 5G NR по сравнению с ограниченными ресурсами микроволнового спектра волн ниже 6 ГГц, которые в настоящее время используются традиционными беспроводными системами и существующими сетями 4G LTE. Очевидно, что используемый микроволновый спектр волн ниже 6 ГГц для существующих беспроводных систем недостаточен для достижения гигабитных скоростей передачи данных. Значительное повышение скорости передачи данных было теоретически подтверждено теоремой Котельникова [2], показывающей, что пропускная способность увеличивается линейно с ростом полосы пропускания. В частности, доступный радиоресурс полос mmWave 71–76 и 81–86 ГГц (известный также, как E-band) [3] значительно превосходит суммарно радиоресурс всех других полос лицензированного спектра. В соответствии с 3GPP Release 16 [3] для раннего развертывания сетей 5G с технологией mmWave были выбраны диапазоны частот 24,25–29,5 и 37–43,5 ГГц. Кроме того, Международным союзом электросвязи (МСЭ) и консорциумом 3GPP был разработан план исследований [4] для стандартов 5G, охватывающих две полосы частот миллиметрового диапазона, которые будут использоваться для коммерческих нужд, а именно: 40 и 100 ГГц. Первоначально еще в 1979 г. для поддержки высокопроизводительной двухточечной связи МСЭ были выделены диапазоны частот E-band — 71–76 и 81–86 ГГц, призванные обеспечить полнодуплексное подключение со скоростью передачи данных 1 Гбит/с и выше. В 2015 г. Федеральная комиссия по связи США предложила гибкие правила обслуживания для полос частот mmWave, включая 28, 37 и 39 ГГц для лицензированного

использования и 64–71 ГГц для нелицензированного использования. Компании AT&T и Verizon для своих начальных развертываний сетей 5G сфокусированы на использовании полос частот 39 и 28 ГГц соответственно.

2 Обзор задач при использовании подвижных точек доступа на беспилотных летательных аппаратах

Перспективным решением для обеспечения пользователей беспроводных сетей широкополосным доступом в местах отсутствия стационарных сотовых базовых станций стало приобретающее в последние годы популярность применение передвижных базовых станций. Так, использование БПЛА, или дронов, и размещения на них точек доступа (ТД) для предоставления услуг связи как в лицензированном (LTE, WiFi), так и в нелицензируемом (mmWave) диапазоне частот обусловлено невероятной гибкостью и скоростью развертывания такой сети. Беспилотные летательные аппараты, оснащенные внешними направленными антennами с возможностью усиления сигнала, выступают не только как средства для инспекций, поиска и фиксации чрезвычайных ситуаций при спасательных операциях в удаленных местах, но и являются частью сетевой транспортной инфраструктуры и служат доступом к услугам передачи данных для пользователей на поверхности земли при проведении всевозможных мероприятий на больших открытых площадках [5–7].

Особенность передачи данных на базе технологии mmWave с помощью БПЛА заключается в необходимости наличия прямой видимости (Line-of-Sight, LoS) между ТД и пользователем. Это жесткое требование компенсируется подвижностью БПЛА, когда обеспечивается адаптивная навигация относительно пользователей, что позволяет значительно повысить эффективность передачи и качество обслуживания [8–11]. В статьях рассмотрен ряд решений для эффективного размещения БПЛА, обеспечивающих оптимизацию услуг передачи данных конечным пользователям.

Число исследований, посвященных применению БПЛА для поддержки беспроводных сетей 5G на базе технологии mmWave, превышает сотню. Благодаря возможности зависания над поверхностью земли, достаточной гибкости и простоте развертывания, высокой маневренности, быстрой реконфигурации сети, БПЛА могут быть размещены практически в любых местах для выполнения функций воздушной ретрансляции на транзитных участках сети в качестве базовой станции или в качестве ТД. При этом, в отличие от БПЛА с фиксированным крылом, у которых положение в пространстве динамически изменяется во времени, положение беспилотных летательных винтокрылых аппаратов относительно фиксированно и статично. В этой связи исследование пространственной конфигурации сети БПЛА стало одной из важных проблем при проектировании беспроводных сетей. Для этого необходимо определять и уметь оптимизировать пространственное положение и траектории БПЛА в соответствии с требованиями

ями к производительности системы. Еще одной ключевой проблемой является возможность эффективного контроля развертывания нескольких БПЛА, или так называемая кластеризация БПЛА и формирование роя БПЛА. Среди последних исследований пространственной конфигурации БПЛА с точки зрения оптимизации положения и траектории известны работы [12–14], вопросы управления развертыванием БПЛА освещены в [15, 16].

Основными преимуществами использования БПЛА с функцией сбора энергии и возможностью хранения данных (кэширования) являются обеспечение непосредственной передачи кэшированного популярного контента в направлении наземных мобильных терминалов и продление операционной работы БПЛА путем сбора энергии из окружающей среды (ветер, солнечная энергия). Первое свойство дополнительно позволяет снизить перегрузки на транзитных участках сети. Однако перерывы в поступлении энергии в силу непостоянности источника и неопределенность кэширования создают дополнительные проблемы для надежного подключения и повсеместного покрытия такими сетями. Вопрос координации и взаимодействия между БПЛА и наземными базовыми станциями описанного сценария рассмотрен в [15]. Кроме того, в данной работе предложена кластеризация БПЛА, направленная на разгрузку трафика в заданном ограниченном пространстве.

Кэширование популярного контента на периферии сети (*edge computing*) в периоды непиковой нагрузки представляется многообещающим решением как для уменьшения нагрузки транзитных соединений в опорных сетях, так и для повышения пропускной способности каналов и улучшения качества восприятия услуг (*Quality of Experience*, QoE) пользователями. Мотивированные этими преимуществами, авторы [12] исследовали проблему развертывания БПЛА с функцией кэширования, чтобы оптимизировать показатели качества восприятия пользователя QoE в облачной сети радиодоступа (*cloud radio access network*, CRAN) путем упреждающей загрузки и кэширования популярного контента в непиковые часы. Для анализа авторами выбрана метрика QoE, характеризующаяся позицией «человек в контуре», которая фиксирует скорость передачи данных, задержку и тип устройства каждого пользователя. В этом случае сформулированная задача оптимизации заключается в минимизации мощности передачи БПЛА при удовлетворении метрики QoE для каждого пользователя. Для решения этой проблемы предложен алгоритм прогнозирования с использованием концепции эхо-сетей (*echo state network*, ESN) для генерации функции распределения моментов времени запроса контента и модели мобильности каждого наземного пользователя, а для навигации БПЛА используется метод k-средних.

Для задач патрулирования, возникающих в приложениях интернета вещей, например при сборе данных с датчиков контроля контура системы безопасности, применяются модели движения БПЛА, не зависящие от перемещения пользователей в области предоставления связи. Модели патрулирования, как правило, представляют собой замкнутые траектории движения, цель которых состоит в поддержании равномерного распределения воздушных базовых станций по

рассматриваемой области. В [17] исследовано семейство траекторий, в соответствии с которым мобильность БПЛА позволяет поддерживать однородный охват пользователей при заданной вероятности покрытия, а также получить преимущество с точки зрения энергоэффективности. В частности, показано, что для рассмотренных частных случаев, а именно: радиальной спиралевидной навигации и кольцевой овальной навигации, — среднее время замирания сигнала (average fade duration, AFD) уменьшается примерно на два порядка по сравнению со статическим случаем расположения БПЛА, например, в соответствии с точечным пуассоновским пространственным процессом.

Еще одно интересное приложение кластеризации дронов представлено в статье [11], посвященной стратегиям развертывания БПЛА в сельской местности. Авторы проводят исследование задачи оптимизации с учетом энергоэффективности и без него и сравнивают скорость выполнения алгоритмов. Системная модель включает малоподвижных, почти статичных пользователей, размещенных внутри прямоугольной области для трех типов пуассоновского точечного процесса — однородного, неоднородного и кластерного. Для навигации БПЛА используются 4 алгоритма: геометрическая релаксация, метод k-средних (k-means), метод k-средних с учетом энергоэффективности и развертывание по сетке. Все алгоритмы максимизируют вероятность покрытия (англ. coverage probability), представляющую собой долю пользователей в области предоставления связи, для которых соединение обеспечивает точки доступа на БПЛА. Ограничением решаемой алгоритмически задачи оптимизации размещения БПЛА служит запрет пересечения зон покрытия ТД, что обеспечивает снижение интерференции. В работе проведено дополнительное исследование зависимости точности решения задачи от погрешности при определении координат пользователей.

3 Примеры навигации беспилотных летательных аппаратов при групповой мобильности пользователей

Решение упомянутых выше задач размещения ТД на БПЛА с помощью методов кластеризации иллюстрируется ниже при сравнительном анализе двух методов определения позиции БПЛА — метода k-средних [18] и метода роя частиц (Particle Swarm Optimization, PSO) [8, 19]. Особенность состоит в моделировании перемещения пользователей с помощью модели групповой мобильности с опорной точкой (Reference Point Group Mobility, RPGM). Модель предусматривает следование группы за условным лидером с одновременным случайным перемещением каждого отдельного объекта внутри группы в ограниченной области на открытой местности [20] и соответствующее динамическое изменение положений БПЛА для максимизации вероятности покрытия. Подключение пользователей к широкополосной беспроводной сети организовано на базе технологии Wi-Fi (семейство протоколов IEEE 802.11, поддерживающих развертывание беспроводных локальных сетей Wireless Local Area Network,

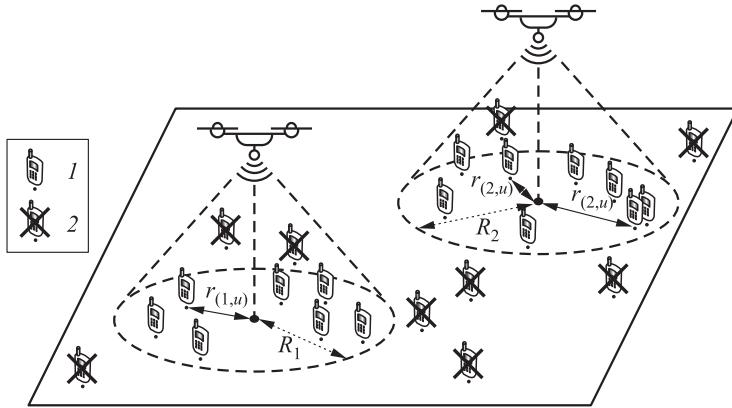


Рис. 1 Сценарий взаимодействия пользователей сети БПЛА: 1 — пользователь в зоне действия БПЛА на высоте h_u ; 2 — пользователь вне зоны действия БПЛА

WLAN) [21], в которой в качестве ТД выступают несколько БПЛА с направленной антенной, обеспечивающих беспроводное радиопокрытие данной области. Физические параметры радиосреды и приемопередающих устройств взяты из [8]. Показателем качества для оценки вероятности покрытия выбрано отношение сигнал/шум (signal-to-noise ratio, SNR), которое должно превосходить заданный стандартами порог $\bar{\gamma}$.

Исследован сценарий взаимодействия объектов беспроводной сети на БПЛА, соответствующий обеспечению связью пользователей при проведении мероприятия на открытой местности без крупных массивных сооружений (здания, деревья, опоры линий электропередачи), так называемый сценарий «концерт» [8]. На рис. 1 показано множество пользователей \mathcal{U} , число которых фиксировано и равно $|\mathcal{U}| = U$, и N_d устройств БПЛА на фиксированной высоте, при этом радиопокрытие каждого БПЛА определяется мощностью излучающей антенны и характеризуется радиусом R_d , $d = 1, \dots, N_d$ [10].

Считаем, что пользователь $u \in \mathcal{U}$ может установить по радиоканалу соединение с ТД на БПЛА d , если $r_{(d,u)} \leq R_d$, где $r_{(d,u)}$ — расстояние между устройством пользователя и центром зоны покрытия ТД на БПЛА. Вероятность покрытия вычисляется как доля пользователей в области покрытия сети на БПЛА:

$$p_c = \frac{1}{U} \sum_{u \in \mathcal{U}} \mathbb{1} \{ \gamma_{(d,u)} \geq \bar{\gamma} \}, \quad d = 1, \dots, N_d,$$

где $\gamma_{(d,u)}$ — отношение сигнал/шум на устройстве пользователя u , принимающего сигнал от БПЛА d , $u \in \mathcal{U}$, $d = 1, \dots, N_d$; $\bar{\gamma}$ — заданный порог отношения сигнал/шум, определяющий качество принимаемого сигнала.

Позиции БПЛА в области представления связи определяются как решение следующей задачи оптимизации:

$$\sum_{d=1}^{N_d} \sum_{u=1}^U 1 \{ r_{(d,u)} \leq R_d \} \rightarrow \max_{(x,y)}$$

при условии $r_{(i,u)} + r_{(j,u)} > R_i + R_j$, $i \neq j$, $j = 1, \dots, N_d$. Здесь $\mathbf{x} = (x_1, \dots, x_{N_d})$, $\mathbf{y} = (y_1, \dots, y_{N_d})$; (x_d, y_d) — координаты центра зоны покрытия ТД на БПЛА d , $d = 1, \dots, N_d$.

Согласно методу k-средних на первом шаге пользователи в зависимости от местоположения разбиваются на кластеры по числу БПЛА. Следующий шаг — построение диаграммы Вороного [22], в результате чего область предоставления связи разделяется на несколько подобластей, в каждой из которых размещается БПЛА. Согласно методу роя частиц изменение позиции подвижной точки доступа подчиняется принципу наилучшего найденного в пространстве положения БПЛА с точки зрения вероятности покрытия.

Исследование проведено для двух случаев разбиения пользователей на группы: (1) группы одинакового размера; (2) число пользователей в группе имеет распределение Ципфа с параметром 0,5. Оба метода использованы для адаптивной навигации, предусматривающей обновление позиций БПЛА через фиксированные интервалы времени t_{upd} . В пользу применения адаптивной навигации говорят результаты для усредненной по 300 с моделирования доле пользователей, находящихся в зоне покрытия БПЛА, приведенные на рис. 2 для групп одинакового размера. Здесь с точки зрения вероятности покрытия метод роя частиц дает выигрыш до 60%, а метод k-средних — до 50% по сравнению с движением БПЛА без учета времени обновления t_{upd} , например при патрулировании по заданным траекториям.

Из рис. 2 также видно, что метод роя частиц дает более высокую вероятность покрытия при сравнимых показателях интерференции. Последнее можно оценить по доле пользователей, находящихся в зоне действия нескольких БПЛА и испытывающих вследствие этого интерференцию, приводящую к снижению качества связи.

Отдельной задачей является выбор длительности интервала t_{upd} . Рисунок 3, а иллюстрирует снижение вероятности покрытия для случая, когда обновление координаты опорной точ-

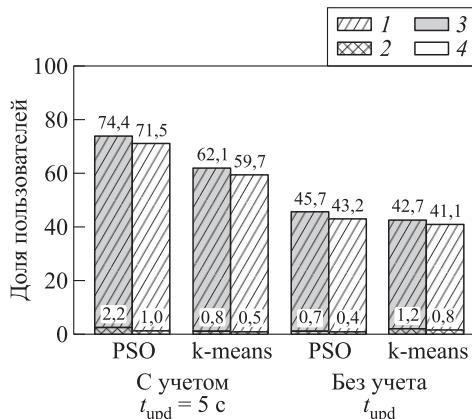


Рис. 2 Средняя доля пользователей в зоне покрытия БПЛА: 1 — пользователи в зоне действия одного БПЛА; 2 — пользователи в зоне действия более одного БПЛА; 3 — число пользователей в группе одинаковое; 4 — число пользователей в группе распределено по Ципфу

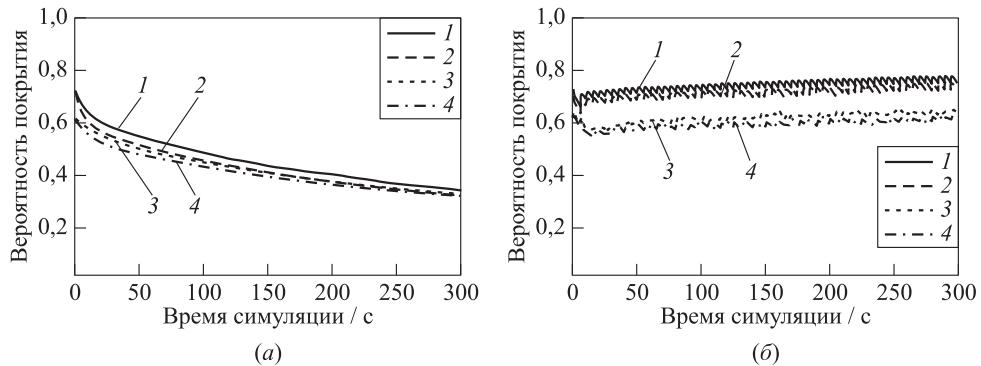


Рис. 3 Вероятность покрытия p_c без учета t_{upd} (а) и с $t_{upd} = 5$ с (б): 1 — PSO, равные группы; 2 — PSO, неравные группы; 3 — k-means, равные группы; 4 — k-means, неравные группы

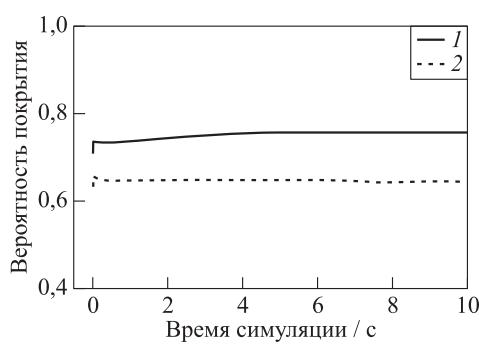


Рис. 4 Вероятность покрытия при $t_{upd} = 0,01$ с: 1 — PSO; 2 — k-means

ны исследована в [17]. В рассмотренном случае для несущей частоты $f_c = 28$ ГГц длина волны $\lambda/2 = 5$ см, что при относительной скорости до 3,6 м/с дает значение $t_{upd} = 10$ мс. Однако, как видно из рис. 4, где приведены результаты моделирования для случая групп равного размера, выигрыш в вероятности покрытия не превышает 4%, что не сравнимо с существенно увеличивающимися затратами на пересчет координат БПЛА и навигацию.

4 Заключение

Анализ статистических характеристик качества покрытия воздушно-наземной беспроводной сетью области моделирования для сценария «концерт» с заданным порогом отношения сигнал/шум показал, что с точки зрения вероятности по-

ки группы проходит слишком часто, что не позволяет БПЛА достичь целевой точки, полученной в результате решения задачи оптимизации. На рис. 3, б корректный выбор длины интервала t_{upd} ведет к росту вероятности покрытия, при этом множественные пики на графике соответствуют значениям вероятности покрытия непосредственно после достижения БПЛА новой позиции, полученной при применении соответствующего алгоритма.

Зависимость интервала t_{upd} от относительной скорости перемещения пользователя и БПЛА и от длины вол-

крытия метод роя частиц для адаптивной навигации позволяет достичь более высоких значений вероятности покрытия, чем метод k-средних. Интересной задачей в дальнейшем видится исследование среднего времени замирания сигнала, представляющего собой среднее время, которое пользователь остается без связи из-за недостаточного качества радиоканала до ТД, обслуживающей соединение пользователя. Авторы планируют в дальнейшем на базе разработанного симулятора расширить сценарии взаимодействия устройств, уделив внимание проблеме построения эффективной зоны покрытия в воздушных беспроводных сетях на БПЛА с учетом расположения одной и нескольких базовых станций, пространственного распределения пользователей внутри зоны покрытия сети и управления распределением поступающей нагрузки от пользователей.

Литература

1. *Zhang L., Zhao H., Hou S., et al.* A survey on 5G millimeter wave communications for UAV-assisted wireless networks // IEEE Access, 2019. Vol. 7. P. 117460–117504.
2. Котельников В. А. О пропускной способности эфира и проволоки в электросвязи: Мат-лы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности. — М.: Всесоюзный энергетический комитет, 1933. (Репринт статьи в журнале УФН, 2006. Т. 176. Вып. 7. С. 762–770.)
3. Technical Specifications and Technical Reports for a 5G based 3GPP system. Specification 21.916.
4. 5G Spectrum Recommendations, 5G Amer. — Bellevue, WA, USA, April 2017.
5. Guillen-Perez A., Sanchez-Iborra R., Cano M. D., Sanchez-Aarnoutse J. C., Garcia-Haro J. WiFi networks on drones // 8th ITU Kaleidoscope Academic Conference: ICTs for a Sustainable World Proceedings. — Bangkok, Thailand, 2016. P. 1–8.
6. Mozaffari M., Saad W., Bennis M., Debbah M. Communications and control for wireless drone-based antenna array // IEEE T. Commun., 2019. Vol. 67. Iss. 1. P. 820–834. doi: 10.1109/TCOMM.2018.2871453.
7. Amer R., Saad W., Marchetti N. Mobility in the sky: Performance and mobility analysis for cellular-connected UAVs // IEEE T. Commun., 2020. Vol. 68. Iss. 5. P. 3229–3246. doi: 10.1109/TCOMM.2020.2973629.
8. Hartigan J., Wong M. A. Algorithm AS 136: A k-means clustering algorithm // J. R. Stat. Soc. C Appl., 1979. Vol. 28. Iss. 1. P. 100–108.
9. Kennedy J., Eberhart R. Particle swarm optimization // IEEE Conference (International) on Neural Networks IV Proceedings. — IEEE, 1995. P. 1942–1948.
10. Zeng Y., Zhang R., Lim T. J. Wireless communications with unmanned aerial vehicles: Opportunities and challenges // IEEE Commun. Mag., 2016. Vol. 54. Iss. 5. P. 36–42.
11. Sun J., Masouros C. Deployment strategies of multiple aerial BSs for user coverage and power efficiency maximization // IEEE T. Commun., 2019. Vol. 67. Iss. 4. P. 2981–2994. doi: 10.1109/TCOMM.2018.2889460.
12. Chen M., Mozaffari M., Saad W., Yin C., Debbah M., Hong C. S. Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized

- quality-of-experience // IEEE J. Sel. Area. Comm., 2017. Vol. 35. Iss. 5. P. 1046–1061.
- 13. Gapeyenko M., Bor-Yaliniz I., Andreev S., Yanikomeroglu H., Koucheryavy Y. Effects of blockage in deploying mmWave drone base stations for 5G Networks and beyond // IEEE Conference (International) on Communications Workshops Proceedings. — IEEE, 2018. Art. ID: 8403671. 6 p. doi: 10.1109/ICCW.2018.8403671.
 - 14. Ghazzai H., Ghorbel M. B., Kassler A., Hossain M. J. Trajectory optimization for cooperative dual-band UAV swarms // IEEE Global Communications Conference Proceedings. — IEEE, 2018. P. 1–7.
 - 15. Wu H., Tao X., Zhang N., Shen X. Cooperative UAV cluster-assisted terrestrial cellular networks for ubiquitous coverage // IEEE J. Sel. Area. Comm., 2018. Vol. 36. Iss. 9. P. 2045–2058. doi: 10.1109/JSAC.2018.2864418.
 - 16. Khosravi Z., Gerasimenko M., Andreev S., Koucheryavy Y. Performance evaluation of UAV-assisted mmWave operation in mobility-enabled urban deployments // 41st Conference (International) on Telecommunications and Signal Processing Proceedings. — IEEE, 2018. P. 150–153. doi: 10.1109/TSP.2018.8441321.
 - 17. Enayati S., Saeedi H., Pishro-Nik H., Yanikomeroglu H. Moving aerial base station networks: a stochastic geometry analysis and design perspective // IEEE T. Wirel. Commun., 2019. Vol. 18. Iss. 6. P. 2977–2988.
 - 18. Tafintsev N., Gerasimenko M., Molchanov D., Akdeniz M., Yeh S., Himayat N., Andreev S., Koucheryavy Y., Valkama M. Improved network coverage with adaptive navigation of mmWave-based drone-cells // IEEE Global Communications Conference Proceedings. — IEEE, 2018. P. 1–7. doi: 10.1109/GLOCOMW.2018.8644097.
 - 19. Xu D., Tian Y. A comprehensive survey of clustering algorithms // Annals Data Science, 2015. Vol. 2. Iss. 2. P. 165–193.
 - 20. Kalantari E., Bor-Yaliniz I., Yongacoglu A., Yanikomeroglu H. User association and bandwidth allocation for terrestrial and aerial base stations with backhaul considerations // 28th Annual Symposium (International) on Personal, Indoor, and Mobile Radio Communications Proceedings. — IEEE, 2017. P. 1–6.
 - 21. Zolanvari M., Jain R., Salman T. Potential data link candidates for civilian unmanned aircraft systems: A survey // IEEE Commun. Surv. Tut., 2020. Vol. 22. Iss. 1. P. 292–319.
 - 22. Preparata F., Shamos M. Computational geometry: An introduction. — Berlin-Heidelberg: Springer-Verlag, 1985. 390 p.

Поступила в редакцию 12.09.20

APPLICATION OF CLUSTERING IN DEPLOYMENT OF MOBILE ACCESS POINTS IN AIR–GROUND WIRELESS NETWORKS

E. G. Medvedeva^{1,2}, E. M. Khayrov¹, N. A. Polyakov¹, and Yu. V. Gaidamaka^{1,2}

¹Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

²Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper provides an overview of tasks that arise in wireless networks with mobile base stations located on unmanned aerial vehicles (UAV). The authors choose two methods of adaptive navigation based on user clustering for a comparative analysis of effectiveness of network deployment — the k-means method and the particle swarm method. The solution to the optimization problem is the positions of UAV that maximize the probability of coverage in the communication provision area with restrictions on interference from neighboring base stations. The application of the methods is illustrated for the scenario of a concert event which is defined as a case for providing communication between participants of a mass event in an open area.

Keywords: UAV; air-ground network; aerial-terrestrial communication; particle swarm; k-means; coverage probability

DOI: 10.14357/08696527200403

Acknowledgments

The publication was prepared with the support of the “RUDN University Program 5-100” and funded by the RFBR according to the research projects No. 18-07-00576 and No. 20-07-01064.

References

1. Zhang, L., H. Zhao, S. Hou, *et al.* 2019. A survey on 5G millimeter wave communications for UAV-assisted wireless networks. *IEEE Access* 7:117460–117504.
2. Kotel’nikov, V. A. 2006. On the transmission capacity of ether and wire in electric communications. *Phys.-Usp.* 49(7):736–744.
3. TS 21.916. 2018. Technical Specifications and Technical Reports for a 5G based 3GPP system. Available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3441> (accessed September 23, 2020).
4. 5G Spectrum Recommendations, 5 G Amer. April 2017. Bellevue, WA. Available at: https://www.5gamerica.org/wp-content/uploads/2019/07/5GA_5G-Spectrum_Recommendations_2017_FINAL.pdf (accessed September 23, 2020).

5. Guillen-Perez, A., R. Sanchez-Iborra, M. D. Cano, J. C. Sanchez-Aarnoutse, and J. Garcia-Haro. 2016. Wifi networks on drones. *8th ITU Kaleidoscope Academic Conference: ICTs for a Sustainable World Proceedings*. Bangkok. 1–8.
6. Mozaffari, M., W. Saad, M. Bennis, and M. Debbah. 2019. Communications and control for wireless drone-based antenna array. *IEEE T. Commun.* 67(1):820–834. doi: 10.1109/TCOMM.2018.2871453.
7. Amer, R., W. Saad, and N. Marchetti. 2020. Mobility in the sky: Performance and mobility analysis for cellular-connected UAVs. *IEEE T. Commun.* 68(5):3229–3246. doi: 10.1109/TCOMM.2020.2973629.
8. Hartigan, J., and M. A. Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. C Appl.* 28(1):100–108.
9. Kennedy, J., and R. Eberhart. 1995. Particle swarm optimization. *IEEE Conference (International) on Neural Networks IV Proceedings*. IEEE. 1942–1948.
10. Zeng, Y., R. Zhang, and T. J. Lim. 2016. Wireless communications with unmanned aerial vehicles: Opportunities and challenges. *IEEE Commun. Mag.* 54(5):36–42.
11. Sun, J., and C. Masouros. 2019. Deployment strategies of multiple aerial BSs for user coverage and power efficiency maximization. *IEEE T. Commun.* 67(4):2981–2994. doi: 10.1109/TCOMM.2018.2889460.
12. Chen, M., M. Mozaffari, W. Saad, C. Yin, M. Debbah, and C. S. Hong. 2017. Caching in the sky: Proactive deployment of cache-enabled unmanned aerial vehicles for optimized quality-of-experience. *IEEE J. Sel. Area. Comm.* 35(5):1046–1061.
13. Gapeyenko, M., I. Bor-Yaliniz, S. Andreev, H. Yanikomeroglu, and Y. Koucheryavy. 2018. Effects of blockage in deploying mmWave drone base stations for 5G networks and beyond. *IEEE Conference (International) on Communications Workshops Proceedings*. IEEE. Art. ID: 8403671. 6 p. doi: 10.1109/ICCW.2018.8403671.
14. Ghazzai, H., M. B. Ghorbel, A. Kassler, and M. J. Hossain. 2018. Trajectory optimization for cooperative dual-band UAV swarms. *IEEE Global Communications Conference Proceedings*. IEEE. 1–7.
15. Wu, H., X. Tao, N. Zhang, and X. Shen. 2018. Cooperative UAV cluster-assisted terrestrial cellular networks for ubiquitous coverage. *IEEE J. Sel. Area. Comm.* 36(9):2045–2058. doi: 10.1109/JSAC.2018.2864418.
16. Khosravi, Z., M. Gerasimenko, S. Andreev, and Y. Koucheryavy. 2018. Performance evaluation of UAV-assisted mmWave operation in mobility-enabled urban deployments. *41st Conference (International) on Telecommunications and Signal Processing Proceedings*. IEEE. 150–153. doi: 10.1109/TSP.2018.8441321.
17. Enayati, S., H. Saeedi, H. Pishro-Nik, and H. Yanikomeroglu. 2019. Moving aerial base station networks: A stochastic geometry analysis and design perspective. *IEEE T. Wirel. Commun.* 18(6): 2977–2988.
18. Tafintsev, N., M. Gerasimenko, D. Moltchanov, M. Akdeniz, S. Yeh, N. Himayat, S. Andreev, Y. Koucheryavy, and M. Valkama. 2018. Improved network coverage with adaptive navigation of mmWave-based drone-cells. *IEEE Global Communications Conference Proceedings*. IEEE. 1–7. doi: 10.1109/GLOCOMW.2018.8644097.
19. Xu, D., and Y. Tian. 2015. A comprehensive survey of clustering algorithms. *Annals Data Science* 2(2):165–193.
20. Kalantari, E., I. Bor-Yaliniz, A. Yongacoglu, and H. Yanikomeroglu. 2017. User association and bandwidth allocation for terrestrial and aerial base stations with

- backhaul considerations. *28th Annual Symposium (International) on Personal, Indoor, and Mobile Radio Communications Proceedings*. IEEE. 1–6.
21. Zolanvari, M., R. Jain, and T. Salman. 2020. Potential data link candidates for civilian unmanned aircraft systems: A survey. *IEEE Commun. Surv. Tut.* 22(1):292–319.
 22. Preparata, F., and M. Shamos. 1985. *Computational geometry: An introduction*. Berlin–Heidelberg: Springer-Verlag. 390 p.

Received September 12, 2020

Contributors

Medvedeva Ekaterina G. (b. 1985) — assistant professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; medvedeva-eg@rudn.ru

Khayrov Emil M. (b. 1997) — Master student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; emil.khayrov@gmail.com

Polyakov Nikita A. (b. 1997) — Master student, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; goto97@mail.ru

Gaidamaka Yuliya V. (b. 1971) — Doctor of Science in physics and mathematics, professor, Department of Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; gaydamaka-yuv@rudn.ru

НЕЧЕТКОЕ УПРАВЛЕНИЕ ГЕТЕРОГЕННЫМ МЫШЛЕНИЕМ АГЕНТОВ ГИБРИДНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ МНОГОАГЕНТНОЙ СИСТЕМЫ*

С. В. Листопад¹, С. Б. Румовская²

Аннотация: Статья посвящена разработке подсистемы нечеткого вывода для управления агентом-фасилитатором (АФ) процессами коллективного гетерогенного мышления агентов гибридной интеллектуальной многоагентной системы. Она позволяет АФ организовать работу агентов системы в соответствии с моделью ромба группового принятия решений С. Кейнера. Модель редуцирует групповую динамику системы на три последовательные фазы: стадию дивергентного мышления, на которойрабатываются альтернативные решения проблемы, стадию бурления, на которой АФ задействует методы, способствующие повышению «взаимопонимания» между агентами, и стадию конвергентного мышления, когда предложенные альтернативы классифицируются, ранжируются и дорабатываются для принятия согласованного решения. Благодаря данному механизму повышается релевантность системы практике коллективного решения проблем экспертами под руководством лица, принимающего решения, при посредничестве фасилитатора. В сочетании с гибридной компонентой интеллектуальной системы и многоагентным подходом, положенным в основу ее архитектуры, моделирование коллективного гетерогенного мышления обеспечивает ей возможность решать практические проблемы без существенного упрощения в условиях динамической среды, ограничивающей время на выработку и принятие решения.

Ключевые слова: гетерогенное мышление; гибридная интеллектуальная многоагентная система; система нечеткого вывода; коллектив экспертов

DOI: 10.14357/08696527200404

1 Введение

Вопросы решения проблем в условиях динамических сред активно исследуются с конца 1970-х гг. [1]. В дополнение к таким НЕ-факторам [2] практических проблем, как неоднородность, недоопределенность, неточность, нечеткость, неполнота, в условиях динамических сред возникают следующие особенности:

* Исследование выполнено при финансовой поддержке РФФИ (проект 18-07-00448а).

¹ Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

² Калининградский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук; Балтийский федеральный университет им. И. Канта, sophiyabr@gmail.com

необходимость принятия нескольких взаимозависимых решений для достижения поставленной цели, изменчивость проблемной среды, принятие решений в реальном времени [3]. Решение проблем в динамической среде характерно для оперативно-диспетчерского управления транспортными потоками, авиаперевозками, энергосетями. Исследования [4–6] показали отрицательное влияние ограничения времени принятия решений на их качество, продемонстрировав существенное упрощение эвристик, которыми руководствуется лицо, принимающее решения. В этой связи актуальна разработка интеллектуальных систем для подготовки решений проблем в динамических средах без их существенного упрощения. Для этого предложены гибридные интеллектуальные многоагентные системы гетерогенного мышления (ГИМСГМ) [7]. Они интегрируют гибридный интеллектуальный подход [8], парадигму многоагентного программирования [9] и моделирование процессов гетерогенного мышления [10]. Настоящая работа посвящена построению системы нечеткого вывода АФ для организации процесса коллективного гетерогенного мышления.

2 Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления

Формально ГИМСГМ определяется следующим образом [7]:

$$\text{himsh} = \langle \text{AG}^*, \text{env}, \text{INT}^*, \text{ORG}, \{\text{ht}\} \rangle .$$

Здесь AG^* — множество агентов ГИМСГМ:

$$\text{AG}^* = \left\{ \text{ag}^{\text{dm}}, \text{ag}^{\text{fc}}, \text{ag}^{\text{med}}, \text{ag}^{\text{int}} \right\} \cup \text{AG}^{\text{ex}} \cup \text{AG}^{\text{it}}, \quad (1)$$

где ag^{dm} — агент, принимающий решения (АПР), выполняющий постановку задач агентам-экспертам (АЭ) и принятие окончательного решения, ag^{fc} — АФ, обеспечивающий эффективное взаимодействие АЭ между собой и с АПР, ag^{med} — агент-посредник, который ведет базу данных зарегистрированных в ГИМСГМ агентов, их имен, возможностей и адресов, ag^{int} — интерфейсный агент, обеспечивающий взаимодействие с пользователем системы, AG^{ex} — подмножество АЭ, генерирующих варианты решения проблемы или ее задач, AG^{it} — подмножество агентов интеллектуальных технологий, позволяющих использовать свои методы другим агентам; env — модель внешней среды ГИМСГМ; $\text{INT}^* = \{\text{prot}_{\text{bsc}}, \text{prot}_{\text{ht}}, \text{lang}, \text{dmscl}\}$ — множество элементов структурирования взаимодействий агентов, включающее базовый протокол взаимодействия агентов prot_{bsc} на основе речевых актов [11], обеспечивающий обмен сообщениями в процессе решения проблемы, протокол коллективного гетерогенного мышления prot_{ht} [12], специфицирующий взаимодействие АФ с АЭ и АПР, язык передачи сообщений lang , модель предметной области ont , классификатор ситуаций

коллективного решения проблемы dmscl [7]; ORG — множество архитектур ГИМСГМ; $\{ht\}$ — множество концептуальных моделей макроуровневых процессов в ГИМСГМ, включающее ht — модель процесса коллективного решения проблем методами гетерогенного мышления — ромб группового принятия решений С. Кейнера, К. Толди, С. Фиск, Д. Бергера [10]. Согласно модели С. Кейнера, процесс решения проблемы ГИМСГМ трехстадийный: (1) дивергентное мышление, обеспечивающее генерацию множества альтернатив; (2) бурление, при котором необходимо сгладить конфликты и сблизить точки зрения агентов; (3) конвергентное мышление, позволяющее доработать предложенные ранее альтернативы для получения коллективного решения, релевантного разнообразию моделей экспертов ГИМСГМ.

Агент $ag \in AG$ из формулы (1) описывается выражением

$$ag = \langle id^{ag}, gl^{ag}, ACT^{ag} \rangle ,$$

где id^{ag} — идентификатор агента; gl^{ag} — цель агента; ACT^{ag} — множество действий, реализуемых агентом. Множество действий АФ включает элементы «управление коллективным гетерогенным мышлением» act_{chtc}^{agfc} , «анализ ситуации коллективного решения проблемы» act_{dmsa}^{agfc} и «выбор метода гетерогенного мышления» act_{htmc}^{agfc} (см. разд. 3), т. е.

$$\left\{ act_{dmsa}^{agfc}, act_{htmc}^{agfc} \right\} \subset ACT^{agfc} .$$

Действие агента из множества ACT^{ag} описывается выражением:

$$act^{ag} = \langle met_{act}^{ag}, it_{act}^{ag} \rangle ,$$

где met_{ag} — метод решения проблемы; it_{ag} — интеллектуальная технология, в рамках которой реализован метод met_{ag} .

Таким образом, функция ГИМСГМ описывается выражением:

$$act_{himsh} = \left(\bigcup_{ag \in AG^*} ACT_{ag} \right) \cup act_{col}, \left| \bigcup_{ag \in AG} \bigcup_{act \in ACT^{ag}} it_{act}^{ag} \right| \geq 2 ,$$

где act_{col} — коллективная функция ГИМСГМ, конструируемая агентами динамически; накладываемое ограничение требует, чтобы в составе ГИМСГМ использовались не менее двух интеллектуальных технологий [8].

3 Управление коллективным гетерогенным мышлением

Для организации коллективного гетерогенного мышления используется метод на основе нечеткого вывода Мамдани со следующими лингвистическими пере-

менными: «конфликт», «развитие конфликта», «стадия», «метод», «длительность». Кроме того, используются следующие детерминированные переменные: «номер итерации», «средняя напряженность конфликта».

Детерминированная переменная «номер итерации» $it \in \mathbb{N}_0 = \{0\} \cup \mathbb{N}$ описывает число повторений действий по анализу текущей ситуации и выбору метода, выполняемых АФ при управлении гетерогенным мышлением.

«Средняя напряженность конфликта» на итерации it определяется как скользящее среднее напряженности конфликта за $n = |\text{AG}^{\text{ex}}|$ итераций:

$$\text{macnf}_{\text{himsh}}^{\text{it}} = \begin{cases} \frac{1}{n} \sum_{i=0}^{n-1} \text{cnf}_{\text{himsh}}^{\text{it}-i} & \text{при } it \geq n; \\ 0 & \text{при } it < n. \end{cases} \quad (2)$$

Здесь $\text{cnf}_{\text{himsh}}^{\text{it}-i}$ — напряженность конфликта в ГИМСГМ на итерации $it - i$, описываемая выражением

$$\text{cnf}_{\text{himsh}}^{\text{it}-i} = \sum_{i=1}^n \sum_{j=i+1}^n 2 \text{cnf}_{\text{himsh}}^{\text{it}-i}(\text{ag}_i, \text{ag}_j) (n-2)!(n!)^{-1},$$

где $\text{cnf}_{\text{himsh}}^{\text{it}-i}(\text{ag}_i, \text{ag}_j)$ — напряженность конфликта пары агентов, определяемая на основе оценки предложенных ими решений методами, рассмотренными в [13].

Лингвистическая переменная «длительность», используемая для оценки продолжительности применения выбранного метода коллективного гетерогенного мышления, представляется выражением:

$$durl = \langle \beta_{\text{durl}}, T_{\text{durl}}, U_{\text{durl}}, G_{\text{durl}}, M_{\text{durl}} \rangle, \quad (3)$$

где β_{durl} = «длительность» — наименование лингвистической переменной; $T_{\text{durl}} = \{\text{«малая»}; \text{«большая»}\}$ — терм-множество ее значений, названий нечеткой переменной; $U_{\text{durl}} = \mathbb{N}_0$ — универсум нечетких переменных; $G_{\text{durl}} = \emptyset$ — процедура образования из элементов множества T_{durl} новых термов; $M_{\text{durl}} = \{\mu_{\text{малая}}, \mu_{\text{большая}}\}$ — процедура, ставящая в соответствие каждому терму множества T_{durl} осмыщенное содержание путем формирования нечеткого множества. При формировании нечетких множеств используются сигмоидальные функции

$$\text{sigmf}(u, b, c) = \left(1 + e^{-b(u-c)}\right)^{-1},$$

где u — элемент нечеткого множества; b и c — параметры функции принадлежности. Первоначально заданы следующие параметры функций принадлежности: $\mu_{\text{малая}} = \text{sigmf}(u_{\text{durl}}; -0,2; 10n)$, $\mu_{\text{большая}} = \text{sigmf}(u_{\text{durl}}; 0,5; 10n)$. Параметры

этих и других функций принадлежности, рассмотренных в работе, должны быть уточнены в ходе тестирования системы.

Лингвистическая переменная «конфликт» для оценки напряженности конфликта между агентами ГИМСГМ задается по аналогии с выражением (3):

$$\text{cnfl} = \langle \beta_{\text{cnfl}}, T_{\text{cnfl}}, U_{\text{cnfl}}, G_{\text{cnfl}}, M_{\text{cnfl}} \rangle, \quad (4)$$

где β_{cnfl} = «конфликт»; $T_{\text{cnfl}} = \{\text{«нет}; \text{«слабый}; \text{«умеренный}; \text{«острый}\}$; $U_{\text{cnfl}} = [0; 1]$; $G_{\text{cnfl}} = \emptyset$; $M_{\text{cnfl}} = \{\mu_{\text{нет}}, \mu_{\text{слабый}}, \mu_{\text{умеренный}}, \mu_{\text{острый}}\}$. При формировании нечетких множеств используются колоколообразные функции:

$$\text{bmf}(u, a, b, c) = \left(1 + |u - c|^{2b}|a|^{-2b}\right)^{-1},$$

где u — элемент нечеткого множества; a , b и c — параметры функции принадлежности.

Первоначально заданы следующие параметры функций принадлежности нечетких множеств: $\mu_{\text{нет}} = \text{bmf}(u_{\text{cnfl}}; 0,15; 5; 0)$; $\mu_{\text{слабый}} = \text{bmf}(u_{\text{cnfl}}; 0,15; 5; 0,3)$; $\mu_{\text{умеренный}} = \text{bmf}(u_{\text{cnfl}}; 0,15; 5; 0,6)$; $\mu_{\text{острый}} = \text{bmf}(u_{\text{cnfl}}; 0,25; 5; 1)$.

Лингвистическая переменная «развитие конфликта», описывающая изменение напряженности конфликта во времени, представляется выражением:

$$\text{dcnfl} = \langle \beta_{\text{dcnfl}}, T_{\text{dcnfl}}, U_{\text{dcnfl}}, G_{\text{dcnfl}}, M_{\text{dcnfl}} \rangle,$$

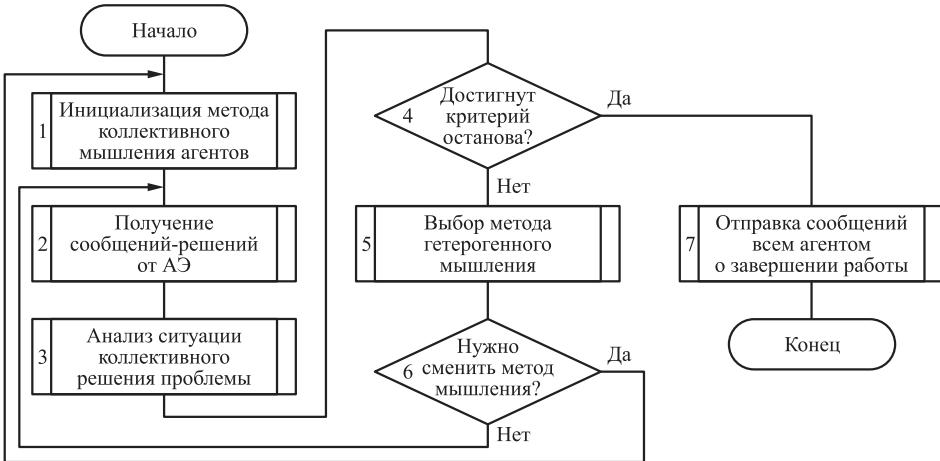
где β_{dcnfl} = «развитие конфликта»; $T_{\text{dcnfl}} = \{\text{«ослабление}; \text{«нет}; \text{«усиление}\}$; $U_{\text{dcnfl}} = [-1; 1]$; $G_{\text{dcnfl}} = \emptyset$; $M_{\text{dcnfl}} = \{\mu_{\text{ослабление}}, \mu_{\text{нет}}, \mu_{\text{усиление}}\}$, $\mu_{\text{ослабление}} = \text{sigmf}(u_{\text{dcnfl}}; -10; -0,3)$, $\mu_{\text{нет}} = \text{bmf}(u_{\text{dcnfl}}; 0,3; 3; 0)$, $\mu_{\text{усиление}} = \text{sigmf}(u_{\text{dcnfl}}; 10; 0,3)$.

Лингвистическая переменная «стадия», сформированная в соответствии с моделью ромба группового принятия решений [10], описывается выражением:

$$\text{stgl} = \langle \beta_{\text{stgl}}, T_{\text{stgl}}, U_{\text{stgl}}, G_{\text{stgl}}, M_{\text{stgl}} \rangle.$$

Здесь β_{stgl} = «стадия»; $T_{\text{stgl}} = \{\text{«дивергентная}; \text{«бурление}; \text{«конвергентная}; \text{«выбор}\}$; $U_{\text{stgl}} = [0; 1]$; $G_{\text{stgl}} = \emptyset$; $M_{\text{stgl}} = \{\mu_{\text{дивергентная}}, \mu_{\text{бурление}}, \mu_{\text{конвергентная}}, \mu_{\text{выбор}}\}$, где $\mu_{\text{дивергентная}} = \text{bmf}(u_{\text{stgl}}; 6^{-1}; 5; 0)$, $\mu_{\text{бурление}} = \text{bmf}(u_{\text{stgl}}; 6^{-1}; 5; 3^{-1})$, $\mu_{\text{конвергентная}} = \text{bmf}(u_{\text{stgl}}; 6^{-1}; 5; 2 \cdot 3^{-1})$, $\mu_{\text{выбор}} = \text{bmf}(u_{\text{stgl}}; 6^{-1}; 5; 1)$.

На основе анализа методов гетерогенного мышления [7] были выделены наиболее перспективные из них для реализации в ГИМСГМ: мозговой штурм с наводящими вопросами (МШНВ); пул мозговой записи (ПМЗ); оспаривание ограничений (ОО); взгляд с чужой точки зрения (ВЧТЗ); возврат от решений



Алгоритм управления коллективным гетерогенным мышлением

к потребностям (ВРП); поиск консенсуса (ПК); голосование экспертов (ГЭ); выбор лицом, принимающим решения, после обсуждения с экспертами (ВЛПРЭ) — и по аналогии с выражением (3) построена лингвистическая переменная «метод»:

$$mtdl = \langle \beta_{mtdl}, T_{mtdl}, U_{mtdl}, G_{mtdl}, M_{mtdl} \rangle.$$

Здесь β_{mtdl} = «метод»; T_{mtdl} = {«МШНВ»; «ПМЗ»; «ОО»; «ВЧТЗ»; «ВРП»; «ПК»; «ГЭ»; «ВЛПРЭ»}; U_{mtdl} = [0; 1]; G_{mtdl} = \emptyset ; M_{mtdl} = $\{\mu_{\text{МШНВ}}, \mu_{\text{ПМЗ}}, \mu_{\text{ОО}}, \mu_{\text{ВЧТЗ}}, \mu_{\text{ВРП}}, \mu_{\text{ПК}}, \mu_{\text{ГЭ}}, \mu_{\text{ВЛПРЭ}}\}$, где $\mu_{\text{МШНВ}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 0)$, $\mu_{\text{ПМЗ}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 7^{-1})$, $\mu_{\text{ОО}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 2 \cdot 7^{-1})$, $\mu_{\text{ВЧТЗ}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 3 \cdot 7^{-1})$, $\mu_{\text{ВРП}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 4 \cdot 7^{-1})$, $\mu_{\text{ПК}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 5 \cdot 7^{-1})$, $\mu_{\text{ГЭ}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 6 \cdot 7^{-1})$, $\mu_{\text{ВЛПРЭ}} = \text{bmf}(u_{mtdl}; 14^{-1}; 5; 1)$.

Алгоритм управления коллективным гетерогенным мышлением на основе рассмотренных лингвистических переменных представлен на рисунке.

Установлены следующие начальные значения переменных: $it = 0$; $u_{stgl}^{it} = 0$; $macnf_{himsh}^{it} = 0$. Алгоритм начинается с инициализации метода коллективного мышления (блок 1): запоминается номер итерации $it^{mtch} = it$ и значение средней напряженности конфликта $macnf_{himsh}^{mtch} = macnf_{himsh}^{it}$. Агент-фасилитатор отправляет АЭ и АПР сообщения, устанавливающие метод коллективного мышления в соответствии со значением лингвистической переменной $mtdl^{it}$

$$mtdl^{it} = \arg \max_{t \in T_{mtdl}} (\mu_t (u_{mtdl}^{it}))$$

и ожидает от них подтверждающие сообщения.

На следующем шаге АФ ожидает от АЭ сообщения, содержащие решения проблемы или ее частей (блок 2). При получении такого сообщения АФ увеличивает значение счетчика итераций $it = it + 1$ и выполняет функцию «анализ ситуации коллективного решения проблемы» act_{dmsa}^{agfc} (блок 3), предназначенную для идентификации текущей стадии коллективного решения проблемы на основе оценки напряженности конфликтов в ГИМСГМ. Для этого значение средней напряженности конфликта (2) присваивается лингвистической переменной «конфликт» (4): $u_{cnfl}^{it} = cnf_{himsht}$, и АФ идентифицирует текущую стадию процесса решения проблемы $stgl^{it}$ в результате нечеткого вывода Мамдани по правилам:

- Правило 1: ЕСЛИ $stgl^{it-1}$ ЕСТЬ «дивергентная»
 И ($cnfl^{it}$ ЕСТЬ «умеренный» ИЛИ $cnfl^{it}$ ЕСТЬ «острый»),
 ТО $stgl^{it}$ ЕСТЬ «бурление». } (5)
- Правило 2: ЕСЛИ $stgl^{it-1}$ ЕСТЬ «бурление»
 И ($cnfl^{it}$ ЕСТЬ «нет» ИЛИ $cnfl^{it}$ ЕСТЬ «слабый»),
 ТО $stgl^{it}$ ЕСТЬ «конвергентная».
- Правило 3: ЕСЛИ $stgl^{it-1}$ ЕСТЬ «конвергентная»
 И $cnfl^{it}$ ЕСТЬ «острый»,
 ТО $stgl^{it}$ ЕСТЬ «бурление».
- Правило 4: ЕСЛИ $stgl^{it-1}$ ЕСТЬ «конвергентная» И $cnfl^{it}$ ЕСТЬ «нет»,
 ТО $stgl^{it}$ ЕСТЬ «выбор».

В блоке 4 оценивается достижение критерия останова: если процесс решения проблемы агентами ГИМСГМ достиг стадии «выбор», т. е. $\arg \max_{t \in T_{stgl}} (\mu_t(u_{stgl}^{it})) = \text{«выбор»}$, выполняется переход к блоку 7 и АФ отправляет АЭ и АПР сообщения о необходимости завершения работы и принятия итогового решения, после чего завершает свою работу. В противном случае выполняется переход к блоку 5.

Для выполнения функции «выбор метода гетерогенного мышления» act_{htmc}^{agfc} (блок 5) должны быть определены значения лингвистических переменных «развитие конфликта», «длительность» и «метод». Значение лингвистической переменной «развитие конфликта» определяется как разница между текущим значением переменной «средняя напряженность конфликта» (2) и значением на момент начала применения метода коллективного гетерогенного мышления $u_{dcnfl}^{it} = macnfl_{himsht} - macnfl_{himsht}^{mtch}$. Значение лингвистической переменной «длительность» — разница между текущим значением переменной «номер итерации» и значением на момент начала применения метода коллективного гетерогенного мышления: $u_{durl}^{it} = it - it^{mtch}$.

Значение лингвистической переменной «метод» определяется в результате нечеткого вывода Мамдани по правилам:

- Правило 1: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «МШНВ»
 И ($dcnfl^{it}$ ЕСТЬ «ослабление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ПМЗ».
- Правило 2: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ПМЗ»
 И ($dcnfl^{it}$ ЕСТЬ «ослабление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ОО».
- Правило 3: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ОО» И ($dcnfl^{it}$ ЕСТЬ «ослабление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ПК»
 И $stgl^{it}$ ЕСТЬ «конвергентная».
- Правило 4: ЕСЛИ $stgl^{it}$ ЕСТЬ «бурление»
 И ($mtdl^{it-1}$ ЕСТЬ «МШНВ»
 ИЛИ $mtdl^{it-1}$ ЕСТЬ «ПМЗ» ИЛИ $mtdl^{it-1}$ ЕСТЬ «ОО»
 ИЛИ $mtdl^{it-1}$ ЕСТЬ «ПК» ИЛИ $mtdl^{it-1}$ ЕСТЬ «ГЭ»
 ИЛИ $mtdl^{it-1}$ ЕСТЬ «ВЛПРЭ»),
 ТО $mtdl^{it}$ ЕСТЬ «ВЧТЗ».
- Правило 5: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ВЧТЗ» И ($dcnfl^{it}$ ЕСТЬ «усиление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ВРП».
- Правило 6: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ВРП» И $durl$ ЕСТЬ «большая»,
 ТО $mtdl^{it}$ ЕСТЬ «ПК» И $stgl^{it}$ ЕСТЬ «конвергентная».
- Правило 7: ЕСЛИ $stgl^{it}$ ЕСТЬ «конвергентная»
 И ($mtdl^{it-1}$ ЕСТЬ «МШНВ»
 ИЛИ $mtdl^{it-1}$ ЕСТЬ «ПМЗ» ИЛИ $mtdl^{it-1}$ ЕСТЬ «ОО»
 ИЛИ $mtdl^{it-1}$ ЕСТЬ «ВЧТЗ» ИЛИ $mtdl^{it-1}$ ЕСТЬ «ВРП»),
 ТО $mtdl^{it}$ ЕСТЬ «ПК».
- Правило 8: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ПК» И ($dcnfl^{it}$ ЕСТЬ «усиление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ГЭ».
- Правило 9: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ГЭ» И ($dcnfl^{it}$ ЕСТЬ «усиление»
 ИЛИ $durl$ ЕСТЬ «большая»), ТО $mtdl^{it}$ ЕСТЬ «ВЛПРЭ».
- Правило 10: ЕСЛИ $mtdl^{it-1}$ ЕСТЬ «ВЛПРЭ» И $durl$ ЕСТЬ «большая»,
 ТО $stgl^{it}$ ЕСТЬ «выбор».

Если выбранный метод коллективного мышления на текущей итерации отличается от метода, выбранного на предыдущей итерации

$$mtd^{it-1} \neq mtd^{it} = \arg \max_{t \in T_{mtd}} (\mu_t (u_{mtd}^{it})) ,$$

выполняется переход к блоку 1, в противном случае — к блоку 2.

Как показывает анализ правил (5) и (6) при управлении АФ процессом коллективного гетерогенного мышления порядок смены методов не предопределен, а зависит от ситуации решения проблемы в ГИМСГМ, в частности длительности использования метода, интенсивности конфликтов, которая определяется составом АЭ и соотношением их целей. Таким образом, в ГИМСГМ релевантно воспроизводится модель ромба группового принятия решений [10]. Благодаря этому ГИМСГМ в условиях динамической среды вырабатывает для каждой проблемы релевантный ей метод решения без упрощения и идеализации по аналогии с тем, как это делает коллектив экспертов под руководством лица, принимающего решения, при посредничестве фасilitатора.

4 Заключение

Рассмотрены особенности решения проблем в условиях динамических сред, в частности отмечено упрощение ситуации и эвристик лицом, принимающим решения. Показана актуальность разработки интеллектуальных систем, способных решать проблемы без их существенного упрощения в условиях дефицита времени. Рассмотрена модель ГИМСГМ, предлагаемых для этой цели. Разработан алгоритм АФ для управления гетерогенным коллективным мышлением АЭ, реализующий модель ромба группового принятия решений С. Кейнера. В основу данного алгоритма положен метод нечеткого вывода Мамдани. Моделирование гетерогенного мышления агентов позволяет ГИМСГМ вырабатывать рекомендации для оперативно-диспетчерского персонала по решению практических проблем в условиях динамической среды.

Литература

1. Каталевский Д. Ю. Основы имитационного моделирования и системного анализа в управлении. — М.: Дело, РАНХиГС, 2015. 496 с.
2. Нариньянц А. С. Инженерия знаний и НЕ-факторы: краткий обзор-08 // Вопросы искусственного интеллекта, 2008. № 1. С. 61–77.
3. Brehmer B. Dynamic decision making: Human control of complex systems // Acta Psychol., 1992. Vol. 81. Iss. 3. P. 211–241.
4. Svenson O., Maule A. J. Time pressure and stress in human judgment and decision making. — New York, NY, USA: Plenum, 1993. 336 p.
5. Maule J. A., Edland A. C. The effects of time pressure on human judgment and decision making // Decision making: Cognitive models and explanations. — New York, NY, USA: Routledge, 1997. P. 189–204.

6. Kerstholt J. H., Raaijmakers J. G. W. Decision making in dynamic task environments // Decision making: Cognitive models and explanations. — New York, NY, USA: Routledge, 1997. P. 205–217.
7. Колесников А. В., Листопад С. В. Модель гибридной интеллектуальной многоагентной системы гетерогенного мышления для информационной подготовки оперативных решений в региональных электрических сетях // Системы и средства информатики, 2018. Т. 28. № 4. С. 31–41.
8. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
9. Тарасов В. Б. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. — М.: Эдиториал УРСС, 2002. 352 с.
10. Kaner S., Lind L., Toldi C., Fisk S., Beger D. The facilitator's guide to participatory decision-making. — San Francisco, CA, USA: Jossey-Bass, 2011. 368 p.
11. Weerasooriya D., Rao A. S., Ramamohanarao K. Design of a concurrent agent-oriented language // Intelligent agents / Eds. M. J. Wooldridge, N. R. Jennings. — Lecture notes in computer science ser. — Springer, 1994. Vol. 890. P. 386–401.
12. Колесников А. В., Листопад С. В. Протокол гетерогенного мышления гибридной интеллектуальной многоагентной системы для решения проблемы восстановления распределительной электросети // Информатика и её применения, 2019. Т. 13. Вып. 2. С. 76–82.
13. Листопад С. В., Кириков И. А. Метод идентификации конфликтов агентов в гибридных интеллектуальных многоагентных системах // Системы и средства информатики, 2020. Т. 30. № 1. С. 56–65.

Поступила в редакцию 30.08.20

FUZZY CONTROL OF HETEROGENEOUS THINKING OF THE HYBRID INTELLIGENT MULTIAGENT SYSTEM'S AGENTS

S. V. Listopad¹ and S. B. Rumovskaya^{1,2}

¹Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation

²Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation

Abstract: The paper is devoted to the development of the fuzzy inference subsystem for controlling the processes of collective heterogeneous thinking of agents of a hybrid intelligent multiagent system by an agent-facilitator. It allows the agent-facilitator to organize the work of the agents of the system in

accordance with S. Kaner's diamond of participatory decision-making model. The model reduces the group dynamics of the system into three sequential phases: the divergent thinking stage at which alternative solutions to the problem are developed; the groan stage at which the agent-facilitator uses methods that enhance "mutual understanding" between the agents; and the convergent thinking stage when the proposed alternatives are classified, ranked, and refined to make an agreed decision. Thanks to this mechanism, the relevance of the system to the practice of collective problem solving by experts under the guidance of a decision maker, with the mediation of a facilitator, increases. In combination with a hybrid component of an intelligent system and a multiagent approach underlying its architecture, modeling of collective heterogeneous thinking provides it with the ability to solve practical problems without significant simplification in a dynamic environment that limits the time to develop and make a decision.

Keywords: heterogeneous thinking; hybrid intelligent multiagent system; fuzzy inference system; expert team

DOI: 10.14357/08696527200404

Acknowledgments

The reported study was funded by RFBR, project No. 18-07-00448a.

References

1. Katalevskiy, D. Yu. 2015. *Osnovy imitatsionnogo modelirovaniya i sistemnogo analiza v upravlenii* [Fundamentals of simulation and system analysis in management]. Moscow: Delo, RANKhiGS. 496 p.
2. Narinyani, A. S. 2008. Inzheneriya znaniy i NE-faktory: kratkiy obzor-08 [Knowledge engineering and nonfactors: A brief overview-08]. *Voprosy iskusstvennogo intellekta* [Artificial Intelligence Issues] 1:61–77.
3. Brehmer, B. 1992. Dynamic decision making: Human control of complex systems. *Acta Psychol.* 81(3):211–241.
4. Svenson, O., and A. J. Maule. 1993. *Time pressure and stress in human judgment and decision making*. New York, NY: Plenum. 336 p.
5. Maule, J. A., and A. C. Edland. 1997. The effects of time pressure on human judgment and decision making. *Decision making: Cognitive models and explanations*. New York, NY: Routledge. 189–204.
6. Kerstholt, J. H., and J. G. W. Raaijmakers. 1997. Decision making in dynamic task environments. *Decision making: Cognitive models and explanations*. New York, NY: Routledge. 205–217.
7. Kolesnikov, A. V., and S. V. Listopad. 2018. Model' gibridnoy intellektual'noy mnogoagentnoy sistemy geterogenного myshleniya dlya informatsionnoy podgotovki operativnykh resheniy v regional'nykh elektricheskikh setyakh [Model of a hybrid intelligent multiagent system of heterogeneous thinking for preparation of information about operational decisions in a regional power system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):31–41.

8. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, dispute]. Moscow: IPI RAN. 189 p.
9. Tarasov, V. B. 2002. *Ot mnogoagentnykh sistem k intellektual'nym organizatsiyam: filosofiya, psichologiya, informatika* [From multiagent systems to intelligent organizations: Philosophy, psychology, and informatics]. Moscow: Editorial URSS. 352 p.
10. Kaner, S., L. Lind, C. Toldi, S. Fisk, and D. Beger. 2011. *The facilitator's guide to participatory decision-making*. San Francisco, CA: Jossey-Bass. 368 p.
11. Weerasooriya, D., A. S. Rao, and K. Ramamohanarao. 1994. Design of a concurrent agent- oriented language. *Intelligent agents*. Eds. M. J. Wooldridge and N. R. Jennings. Lecture notes in computer science ser. Springer. 890:386–401.
12. Kolesnikov, A. V., and S. V. Listopad. 2019. Protokol geterogenного мышления гибридной интеллектуальной многоагентной системы для решения проблемы восстановления распределительной электросети [Heterogeneous thinking protocol of hybrid intelligent multiagent system for solving distributional power grid recovery problem]. *Informatika i ee Primeneniya — Inform. Appl.* 13(2):76–82.
13. Listopad, S. V., and I. A. Kirikov. 2020. Metod identifikatsii konfliktov agentov v gibridnykh intellektual'nykh mnogoagentnykh sistemakh [Agent conflict identification method in hybrid intelligent multiagent systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):56–65.

Received August 30, 2020

Contributors

Listopad Sergey V. (b. 1984) — Candidate of Science (PhD) in technology, senior scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; ser-list-post@yandex.ru

Rumovskaya Sophiya B. (b. 1985) — Candidate of Science (PhD) in technology, scientist, Kaliningrad Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 5 Gostinaya Str., Kaliningrad 236000, Russian Federation; senior lecturer, Immanuel Kant Baltic Federal University, 14 A. Nevskogo Str., Kaliningrad 236041, Russian Federation; sophiyab@gmail.com

СТАТИСТИЧЕСКИЙ АНАЛИЗ БИОСИГНАЛОВ ПРИ ИЗУЧЕНИИ ЗРИТЕЛЬНОЙ РАБОЧЕЙ ПАМЯТИ ЧЕЛОВЕКА*

A. V. Ерофеева¹, T. B. Захарова²

Аннотация: Работа посвящена электроэнцефалографическому исследованию взаимодействия корковых зон мозга, обеспечивающих удержание зрительной информации в рабочей памяти (РП). Были построены векторные авторегрессионные модели (VAR-модели) сигналов, полученных из связанных с организацией РП зон мозга. Для оценки силы взаимодействия зон по коэффициентам моделей рассчитывалась функция частной направленной коррентности (Partial Directed Coherence, PDC), основанная на причинности Грейнджа. Сравнительный анализ для оценки силы внутрикорковых связей был проведен с помощью парного статистического теста Уилкоксона. Найдена зависимость силы связей от характера выполняемой задачи.

Ключевые слова: ЭЭГ; функциональная связность; зрительная рабочая память; PDC; векторная авторегрессионная модель; тест Уилкоксона; причинность Грейнджа

DOI: 10.14357/08696527200405

1 Введение

Рабочая память — многокомпонентная система, несущая функцию временного хранения информации в активном и доступном для анализа состоянии. Рабочая память принимает участие в мышлении человека, служит связующим звеном между системами восприятия и долговременной памятью, а характеристики РП определяют скорость и объем обработки информации. В составе РП выделяют зрительную РП — подсистему, использующуюся для хранения зрительной информации и для манипулирования ею в ходе выполнения познавательных заданий [1–3].

В настоящее время механизмы РП активно изучаются. Хотя была показана ведущая роль префронтальных отделов коры головного мозга в обеспечении

*Работа выполнена при частичной поддержке РФФИ (проект 18-07-00252) и в соответствии с программой Московского центра фундаментальной и прикладной математики.

¹Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, yeofeyeva@bk.ru

²Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, tvzaharova@mail.ru

функций РП, все чаще исследуются и другие зоны мозга. За последние годы накопилось множество экспериментальных данных, согласно которым при выполнении людьми заданий, требующих РП, активность наблюдается не только в префронтальной, но и в зрительной, теменной, височной коре [4, 5]. Цель данной работы — изучение связности вышеназванных зон мозга и их взаимодействия при удержании информации в зрительной РП.

2 Описание эксперимента

Исследование функциональной связности мозга было проведено при помощи следующего эксперимента. Эксперимент включал в себя две серии испытаний: «Наблюдение» и «Рабочая память». Каждое испытание предполагало наблюдение стимулов — черно-белых изображений прямоугольных решеток с наклоном линий в первой серии и троек небольших черных кружков на белом фоне во второй. Изображения стимулов представлены на рис. 1 и 2.

В «Наблюдении» для испытуемых не предполагалось когнитивной задачи. В «Рабочей памяти» моделировались ситуации сравнения испытуемым новой и удерживаемой в памяти информации. В случайном порядке предъявляли пары совпадающих и несовпадающих стимулов. Испытуемый должен был запомнить первый (эталонный) стимул в паре, сравнить его со вторым стимулом (тестовым) и определить, совпадают ли они. Более подробное описание испытаний представлено на рис. 3 и 4.



Рис. 1 Стимулы первой серии — наклонные решетки

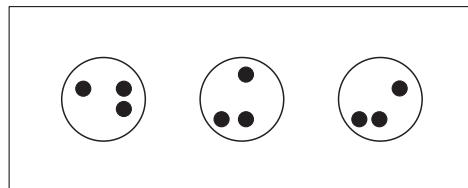


Рис. 2 Стимулы второй серии — паттерны кружков

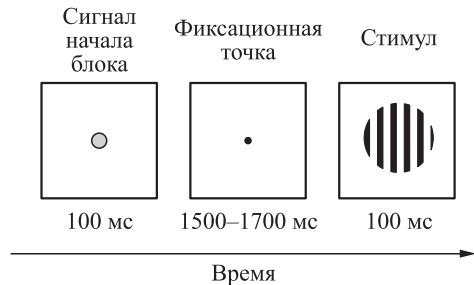


Рис. 3 Испытание в серии «Наблюдение»

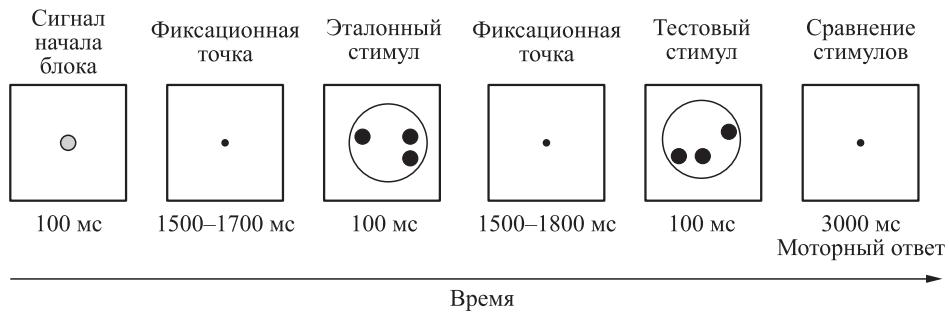


Рис. 4 Испытание в серии «Рабочая память»

3 Моделирование электроэнцефалограммы векторной авторегрессионной моделью

Ключевую роль в исследовании играет VAR-моделирование. На основе характеристик построенной модели будет сделан вывод о степени влияния каналов друг на друга.

Опишем с помощью VAR-модели порядка p многоканальный временной ряд электроэнцефалограммы (ЭЭГ). Пусть дан отрезок оцифрованной M -канальной ЭЭГ длительностью N отсчетов:

$$X(t) = [X_1(t), \dots, X_M(t)]^T .$$

Тогда $X(t)$ имеет вид:

$$X(t) = \sum_{i=1}^p A(i)X(t-i) + E(t) ,$$

где $A(i)$ — матрица размерности $M \times M$; $E(t)$ — ошибка предсказания, вектор белого шума размера M ; p — порядок модели.

Элемент $a_{km}(i)$ матрицы $A(i)$ показывает, какой вклад в текущее значение сигнала $X_k(t)$ вносит сигнал $X_m(t - i)$. Полное влияние канала m на канал k определяется последовательностью коэффициентов $a_{km}(i)$ при $i \in \{1, \dots, p\}$. Если коэффициенты равны нулю для всех i , то считаем, что влияние отсутствует.

Следует отметить, что построение VAR-моделей имеет смысл лишь для стационарных сигналов, которым ЭЭГ-сигнал в общем случае не является. Не-стационарность ЭЭГ, однако, не исключает применения VAR-моделирования: если найдется промежуток, в пределах которого сигнал стационарен, то построенная на нем VAR-модель будет давать адекватные результаты. Сигналы со стационарными промежутками называются локально стационарными.

Кроме того, важен выбор порядка VAR-модели, т. е. числа точек в прошлом, которые будут использоваться для предсказания будущего состояния: VAR-модель с малым числом задержек аппроксимирует ЭЭГ некачественно, модели высокого порядка более детализированы, но часто ненадежны [6].

Для нахождения наилучшего порядка модели может использоваться критерий Акаике (AIC, Akaike's Information Criterion). Для фрагмента многоканальной ЭЭГ строится серия VAR-моделей от 1-го до некоторого порядка P_m , заведомо превосходящего оптимальный. Для каждой модели порядка $p \in \{1, \dots, P_m\}$ определяется ковариационная матрица остатков $\Sigma(p)$ и вычисляется значение функции AIC:

$$\text{AIC}(p) = N \ln(\det \Sigma(p)) + 2pM^2.$$

Считается, что чем меньше значение, тем лучше модель описывает данные. Однако в качестве оптимального порядка модели рекомендуется выбирать либо точку первого локального минимума функции $\text{AIC}(p)$, либо в случае монотонного убывания функции ограничивать максимальный порядок, например, величиной [7]:

$$p_{\max} < \frac{3\sqrt{N}}{M}. \quad (1)$$

4 Расчет частной направленной когерентности

В последние годы популярность в исследовании связности головного мозга набирают методы, основанные на причинности Грейнджера. Причинность Грейнджера (Granger causality) — один из способов исследования влияния временных рядов друг на друга, позволяющий не только находить скрытые взаимодействия и оценивать их силу, но и определять направления их действия. Идея этого метода заключается в построении предсказательных моделей, и если модель при учете предыдущих значений некоторого временного ряда X точнее предсказывает поведение ряда Y , то считается, что X влияет на Y по Грейндже.

При исследовании влияния одного канала на другой может потребоваться знать, какой вклад вносит каждый ритм. Ритмами называют электрические

колебания головного мозга, соответствующие определенному частотному диапазону. Считается, что каждый ритм соответствует некоторому состоянию мозга и отражает процессы, которые происходят в нем. Характеристики ритма могут быть оценены путем вычисления частотных характеристик его диапазона.

В рамках векторной авторегрессионной модели в качестве меры влияния по Грейндджеру одного канала на другой на частоте f может быть использована PDC-функция частной направленной когерентности:

$$P_{ij}(f) = \frac{|A_{ij}(f)|}{\sqrt{\sum_{k=1}^M |A_{kj}(f)|^2}},$$

где $A_{ij}(f)$ — элемент $A(f)$:

$$A_{ij}(f) = \begin{cases} 1 - \sum_{r=1}^p a_{ij}(r)e^{-i2\pi fr}, & i = j; \\ \sum_{r=1}^p a_{ij}(r)e^{-i2\pi fr}, & i \neq j. \end{cases}$$

Элемент $A_{ij}(f)$ матрицы $A(f)$ показывает, какой частотной фильтрации подвергнется сигнал X_j , прежде чем стать частью сигнала X_i . Поэтому величину $P_{ij}(f)$ можно рассматривать как меру частотного направленного влияния канала j на канал i , нормированную на совокупное влияние j на все каналы, включая каналы i и его самого.

5 Обработка данных и построение модели

Из записей ЭЭГ каждого испытуемого исключались фрагменты с физиологическими артефактами и с неверными ответами. Далее выделялись отрезки длиной 900 мс, соответствующие этапу удержания информации в РП, и для каждого слабостационарного 900-миллисекундного отрезка строилась векторная авторегрессионная модель.

В качестве каналов модели в полушариях головного мозга были выбраны симметричные сенсоры, соответствующие затылочному, теменному, височному и лобному отделам коры. Исходя из расположения электродов, пары были разделены на длинные (пары с сенсорами лобного отдела коры) и короткие функциональные связи. Всего образовано 12 пар сенсоров: по три пары длинных и коротких связей в каждом полушарии. На рис. 5 изображены связи, наложенные на схему монтажа электродов шлема ЭЭГ. По данным из выбранных электродов были рассчитаны коэффициенты авторегрессионных моделей.

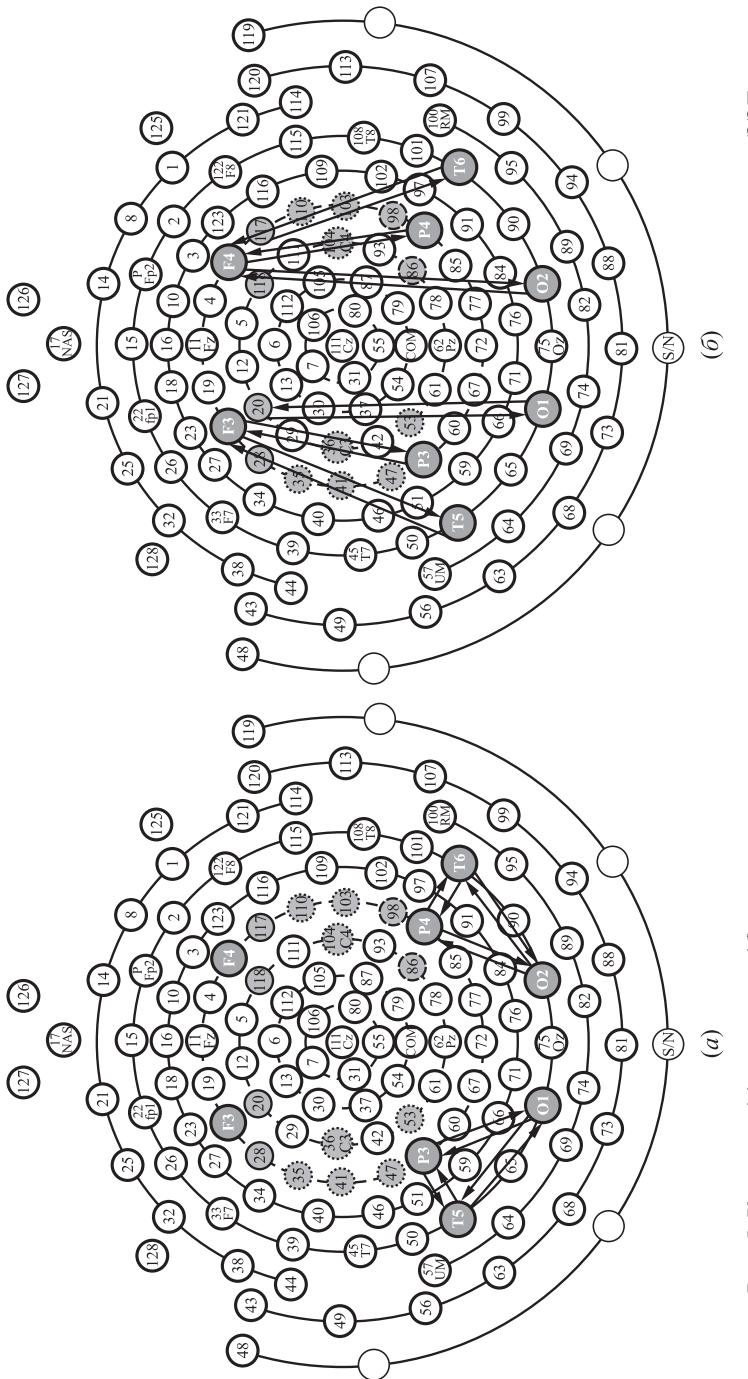


Рис. 5 Корткие (а) и длинные (б) внутрикорковые связи, наложенные на схему монтажа электродов шлема ЭЭГ

Для нахождения оптимального порядка VAR-моделей из каждого эксперимента случайным образом выбирались промежутки, на которых строилась функция AIC(p) для $p \in \{1, \dots, 20\}$. В каждом эксперименте функции монотонно убывали и для всех выбранных p их значения лежали в промежутке от 3 до 7. Учитывая верхнюю оценку (1), отрезки для каждого эксперимента приходилось моделировать VAR-моделями 15-го порядка.

Коэффициенты моделей использовались для вычисления PDC. Оценки рассчитывались для частотного диапазона от 4 до 30 Гц с шагом в 0,5 Гц для всех стационарных 900-миллисекундных сегментов, а затем усреднялись по ним для каждой частоты. Описанная выше процедура оценки была необходима для исключения неодинакового статистического смещения, возникающего при сравнении оценок, полученных для временных эпох разной длительности [8, 9]. Полученные значения функции частной направленной когерентности усреднялись в частотных диапазонах: θ (4–7,5 Гц), α (8–12,5 Гц), β_1 (13–20,5 Гц) и β_2 (21–30 Гц).

Так как значения направленной когерентности не являлись нормально распределенными, статистический анализ их различий для пар сенсоров в задачах наблюдения и удержания в памяти для каждого из типов стимулов (линии и паттерны) проводился с использованием парного критерия Уилкоксона с уровнем значимости $\alpha = 0,05$.

6 Результаты

В работе был проведен сравнительный анализ внутрикорковых связей в задаче зрительной рабочей памяти на наклон линий и расположение паттернов. Связи оценивались на этапе удержания зрительного стимула и сравнивались с корковыми связями при простом наблюдении этих же стимулов.

Статистический анализ проведен с помощью парного теста Уилкоксона с уровнем значимости $\alpha = 0,05$ и с учетом поправки Холма–Бонферрони на множественные сравнения. В рамках каждой серии для каждого полушария проверялось 48 гипотез.

В серии на ориентации линий с учетом поправки Холма–Бонферрони значимые результаты были получены для обоих полушарий. В левом полушарии значимые различия PDC были найдены для пары сенсоров F3 → T5 (нисходящая связь от лобного сенсора к височному) в диапазонах α , β_1 и β_2 с pvalue, равными 0,0003, 0,0001, 0,0001 соответственно, а также для пары F3 → P3 (нисходящая связь от лобного сенсора к теменному) в диапазонах θ и β_2 с pvalue, равными 0,0006 и 0,001.

В правом полушарии значимое различие PDC было найдено для пары сенсоров F4 → T6 (также нисходящая связь от лобного сенсора к височному) в диапазонах α и β_1 с pvalue, равными 0,001 и 0,0007. Полученные статистически значимые связи, наложенные на схему монтажа электродов шлема ЭЭГ, изображены на рис. 6.

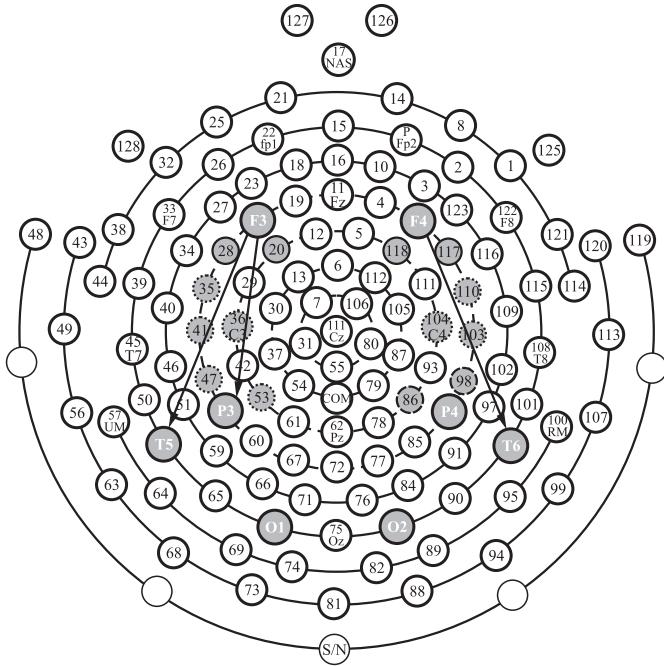


Рис. 6 Внутрикорковые связи со значимыми различиями PDC при наблюдении и удержании информации, наложенные на схему монтажа электродов шлема ЭЭГ

В серии на паттерны не было получено значимых результатов с учетом поправки Холма–Бонферрони как для левого, так и для правого полушарий.

7 Заключение

В данной работе был применен метод аппроксимирования ЭЭГ авторегрессионной моделью. С помощью функции частной направленной когерентности были определены силы влияния каналов ЭЭГ друг на друга. Использование теста Уилкоксона позволило оценить силу статистической связи между каналами ЭЭГ. В задаче зрительной рабочей памяти на ориентации линий найдены значимые различия в уровне нисходящей связи от лобной к височной области в обоих полушариях. Также в задаче на ориентации линий в левом полушарии найдены значимые различия в уровне нисходящей связи от лобной к теменной области. Таким образом, в работе показана зависимость силы длинных связей от характера выполняемой задачи. Полученные результаты могут быть полезны в изучении болезней, для которых характерны различные нарушения рабочей памяти.

Литература

1. Величковский Б. Б., Козловский С. А. Рабочая память человека: фундаментальные исследования и практические приложения // Интеграл, 2013. Т. 68. № 6. С. 38–40.
2. Уточкин И. С., Юрьевич М. А., Булатова М. Е. Зрительная рабочая память: методы, исследования, теории // Российский ж. когнитивной науки, 2016. Т. 3. № 3. С. 58–76.
3. Christophel T. B., Klink P. C., Spitzer B., Roelfsema P. R., Haynes J. D. The distributed nature of working memory // Trends Cogn. Sci., 2017. Vol. 21. Iss. 2. P. 111–124.
4. Ivanitsky A. M., Nikolaev A. R., Ivanitsky G. A. Electroencephalography // Modern techniques in neuroscience research. — Berlin–Heidelberg: Springer-Verlag, 1999. P. 971–995.
5. Baccala L. A., Sameshima K. Partial directed coherence: A new concept in neural structure determination // Biol. Cybern., 2001. Vol. 84. P. 463–474.
6. Курганский А. В. Некоторые вопросы исследования корково-корковых функциональных связей с помощью векторной авторегрессионной модели многоканальной ЭЭГ // Ж. высшей нервной деятельности им. И. П. Павлова, 2010. Т. 60. № 6. С. 740–759.
7. Marple S. L., Jr. Digital spectral analysis with applications. — Englewood Cliffs, NJ, USA: Prentice-Hall, 1987. 218 p.
8. Ding M., Bressler S. L., Yang W., Liang H. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: Data preprocessing, model validation, and variability assessment // Biol. Cybern., 2000. Vol. 83. Iss. 1. P. 35–45.
9. Schlogl A. The electroencephalogram and the adaptive autoregressive model: Theory and applications. — Graz: Technischen Universitat Graz, 2000. PhD Diss. 93 p.

Поступила в редакцию 03.06.20

BIOSIGNAL STATISTICAL ANALYSIS IN THE STUDY OF HUMAN VISUAL WORKING MEMORY

A. V. Erofeeva¹ and T. V. Zakharova^{1,2}

¹Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

²Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The article is devoted to EEG (electroencephalography) studying of connectivity of cortical areas involved in keeping vision information in the

working memory. There was used VAR-modeling for describing signals got from connected with working memory brain zones. Brain connections were estimated the based on Granger Causality Partial Directed Coherence (PDC) and then compared by Wilcoxon signed-rank test. Connection intensity dependence on the executing task was found.

Keywords: EEG; functional connectivity; vision working memory; PDC; VAR-model; Wilcoxon signed-rank test; Granger causality

DOI: 10.14357/08696527200405

Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 18-07-00252). The research was conducted in accordance with the program of the Moscow Center for Fundamental and Applied Mathematics.

References

1. Velichkovsky, B. B., and S. A. Kozlovsky. 2013. Rabochaya pamiat' cheloveka: fundamental'nye issledovaniya i prakticheskie prilozheniya [Human working memory: Fundamental researches and practical applications]. *Integral* [Integral] 68(6):38–40.
2. Utochkin, I. S., M. A. Yurevich, and M. E. Bulatova. 2016. Zritel'naya rabochaya pamiat': metody, issledovaniya, teorii [Vision working memory: Methods, research, theory]. *Rossiyskiy zh. kognitivnoy nauki* [Russ. J. Cognitive Science] 3(3):58–76.
3. Christophel, T. B., P. C. Klink, B. Spitzer, P. R. Roelfsema, and J. D. Haynes. 2017. The distributed nature of working memory. *Trends Cogn. Sci.* 21(2):111–124.
4. Ivanitsky, A. M., A. R. Nikolaev, and G. A. Ivanitsky. 1999. *Electroencephalography. Modern techniques in neuroscience research*. Berlin–Heidelberg: Springer-Verlag. 971–995.
5. Baccala, L. A., and K. Sameshima. 2001. Partial directed coherence: A new concept in neural structure determination. *Biol. Cybern.* 84:463–474.
6. Kurganskii, A. V. 2010. Nekotorye voprosy issledovaniya korkovo-korkovykh funktsional'nykh svyazey s pomoshch'yu vektornoy avtoregressionnoy modeli mnogokanal'noy EEG [Study of cortico-cortical functional connectivity with vector autoregressive model of multichannel EEG]. *Zh. vysshey nervnoy deyatel'nosti I. P. Pavlova* [I. P. Pavlov J. Higher Nervous Activity] 60(6):740–759.
7. Marple, S. L., Jr. 1987. *Digital spectral analysis with applications*. Englewood Cliffs, NJ: Prentice-Hall. 218 p.
8. Ding, M., S. L. Bressler, W. Yang, and H. Liang. 2000. Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: Data preprocessing, model validation, and variability assessment. *Biol. Cybern.* 83(1):35–45.
9. Schlogl, A. 2000. The electroencephalogram and the adaptive autoregressive model: Theory and applications. Graz: Technischen Universitat Graz. PhD Diss. 93 p.

Received June 3, 2020

Contributors

Erofeeva Arina V. (b. 1997) — student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; yerofeyeva@bk.ru

Zakharova Tatiana V. (b. 1962) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; lsa@cs.msu.ru

ПОСТРОЕНИЕ МОДЕЛЕЙ ПРОЦЕССА С ПОМОЩЬЮ ПРОСТЫХ СЕТЕЙ ПЕТРИ*

И. Ю. Терёхина¹, А. А. Грушо², Е. Е. Тимонина³, С. Я. Шоргин⁴

Аннотация: Цель данной работы — исследование класса простых сетей Петри для моделирования рабочего процесса. Показано, что при отсутствии ограничения на уникальность задач в логе построить корректную модель рабочего процесса не всегда удается. Также показано, что если в модели есть переходы, которые не соответствуют ни одной задаче лога рабочего процесса, то в модели рабочего процесса будут нарушены предположения о взаимосвязи между задачами в логе рабочего процесса и между переходами соответствующей сети Петри.

Ключевые слова: сеть Петри; обнаружение процесса; построение модели процесса

DOI: 10.14357/08696527200406

1 Введение

Под термином майнинг (обнаружение) процесса понимается метод построения модели процесса из множества примеров реализации процесса. Задача обнаружения процесса не нова, но все еще остается актуальной. В контексте рабочих процессов данная задача впервые изучалась в работе [1], после чего появилось множество различных подходов к ее решению (см., например, [2–6]).

Основные предположения, которые, как правило, накладываются на рабочий процесс и его лог:

- каждое событие относится к некоторой задаче (шаг рабочего процесса);
- каждое событие относится к некоторой реализации рабочего процесса;
- события упорядочены (притом что задачи могут исполняться параллельно);
- события атомарны.

*Работа частично поддержана РФФИ (проект 18-07-00274).

¹Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, iteroyokhina@cs.msu.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, grusho@yandex.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, eltimon@yandex.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sshorgin@ipiran.ru

В работе [6] описываются техники для обнаружения моделей рабочих процессов в терминах сетей Петри [7] и предложен алгоритм, который может обнаруживать рабочий процесс, представленный его логом. Считается [8], что моделирование в терминах сетей Петри помогает получить дополнительную информацию о структуре и динамическом поведении моделируемой системы.

Цель данной работы — исследование предложенного в работе [6] класса сетей Петри для моделирования рабочего процесса. Показано, что при отсутствии ограничения на уникальность задач в логе построить корректную модель рабочего процесса не всегда удается. Также показано, что если в модели есть переходы, которые не соответствуют никакой задаче в логе, в частности вспомогательные переходы AND-split, AND-join, OR-split, OR-join, то нарушаются теоремы работы [6] о взаимосвязи между задачами в логе рабочего процесса и между переходами соответствующей сети Петри. Тем самым показывая, что алгоритм обнаружения процесса по логу из работы [6] будет работать лишь для узкого класса рабочих процессов, в которых нет повторов задач и условий «ИЛИ».

2 Необходимые свойства модели для ее обнаружения

Под сетью Петри будем понимать тройку $N = (P, T, F)$, где P — конечное множество *мест*; T — конечное множество *переходов*, $P \cap T = \emptyset$; $F \subseteq (P \times T) \cup (T \times P)$ — множество направленных некратных дуг, отношение инцидентности.

Для описания динамического поведения сети Петри вводится определение *разметки* над множеством мест P . Обычно под разметкой понимается мульти множество мест из P [6]. Так, запись $s = \{p_1, p_1, p_2\}$ означала бы, что в сети N есть 3 метки, 2 из них находятся на месте $p_1 \in P$, одна метка стоит на месте $p_2 \in P$. В рамках данной работы будут рассматриваться разметки, в которых на каждом месте сети N может находиться не более одной метки. Таким образом, разметка s — это множество мест из P , на которых есть метка.

Пусть \mathcal{N} — множество всех размеченных сетей Петри, т. е. пар вида (N, s) .

В работе [6] вводится ряд свойств корректности, накладываемых на сеть Петри для того, чтобы модель могла быть обнаружена по логу рабочего процесса. В данной работе будет проведен анализ совместимости свойств:

- (1) надежности;
- (2) запрета последовательного выполнения в сети конструкций синхронизации и выбора;
- (3) полноты лога рабочего процесса для рассматриваемой сети;
- (4) сохранения в сети отношения каузальности между переходами, если между соответствующими задачами в логе это отношение было выполнено.

Введем определения и сформулируем теорему в соответствии с вышеупомянутыми свойствами. Далее будут рассматриваться сети Петри $N = (P, T, F)$ с одним входным и выходным местом такие, что $\forall x, y \in P \cup T, xF^*y$, где F^* —

рефлексивное и транзитивное замыкание отношения F . Элементы из $P \cup T$ будем называть *узлами*. Введем обозначения входных и выходных узлов для некоторого узла x сети N : $(\cdot x) = \{y | (y, x) \in F\}$, $(x \cdot) = \{y | (x, y) \in F\}$.

Динамическое поведение сети Петри описывается правилом срабатывания переходов.

Определение 1 (срабатывание переходов). Пусть $(N = (P, T, F), s)$ — сеть Петри с метками. Переход $t \in T$ называется *активированным*, $(N, s)[t]$, если $(\cdot t) \subseteq s$. Пусть $t \in T$ — активированный переход в сети N . Срабатывание перехода t приводит сеть N с разметкой s к новой разметке $s \cup (t \cdot) \setminus (\cdot t)$ обозначение $(N, s)[t](N, s \cup (t \cdot) \setminus (\cdot t))$.

Определение 2 (достижимые разметки). Пусть (N, s_0) — размеченная сеть Петри из \mathcal{N} . Разметка s называется *достижимой* из разметки s_0 , если существует последовательность активированных переходов, таких что их последовательное срабатывание приведет из разметки s_0 в разметку s . Множество достижимых разметок (N, s_0) обозначается $[N, s_0]$.

Далее будут рассматриваться сети Петри с двумя выделенными местами: $i \in P$ — входное место сети и $o \in P$ — выходное место сети. Предполагается, что в сеть N новая метка может поступать только на входное место и покидать сеть только через выходное место сети.

Рассмотрим пример размеченной сети Петри $(N = (P, T, F), \{i\})$, где $P = \{i, o, p_1, p_2\}$, $T = \{i, o, t_1, t_2\}$, $F = \{(i, t_1), (t_1, p_1), (t_1, p_2), (p_1, t_2), (t_2, o)\}$ (рис. 1). Метка на месте i активирует переход t_1 . Множество входящих мест в переход t_1 : $(\cdot t_1) = \{i\}$, множество выходящих мест $(t_1 \cdot) = \{p_1, p_2\}$

Пусть срабатывает переход t_1 . Стоит отметить, что если в сети Петри есть несколько активированных переходов, то в следующий момент времени может сработать любое подмножество из них [8]. В данной работе предполагается, что в один момент времени может сработать только один из активированных переходов. По правилу срабатывания переходов новая разметка для сети N после срабатывания перехода: $t_1 : s \cup (t \cdot) \setminus (\cdot t) = \{i\} \cup \{p_1, p_2\} \setminus \{i\} = = \{p_1, p_2\}$, становится активированным переход t_2 (рис. 2).

Пусть срабатывает переход t_2 . Множество входящих мест в переход t_2 : $(\cdot t_2) = \{p_1\}$, множество выходящих мест $(t_2 \cdot) = \{o\}$. Новая разметка после срабатывания перехода t_2 по правилу срабатывания переходов: $\{p_1, p_2\} \cup \{o\} \setminus \{p_1\} = = \{p_2, o\}$ (рис. 3).

Множество достижимых разметок для примера на рис. 1: $[N, \{i\}] = = \{\{i\}, \{p_1, p_2\}, \{p_2, o\}\}$.

Свойство 1 определяется следующим образом.

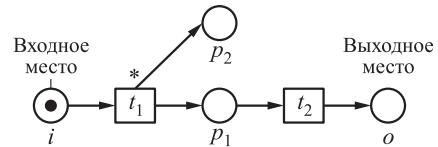


Рис. 1 Размеченная сеть $(N, \{i\})$. Активированный переход отмечен знаком «*»

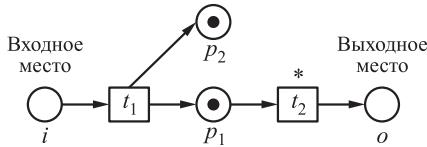


Рис. 2 Размеченная сеть $(N, \{p_1, p_2\})$

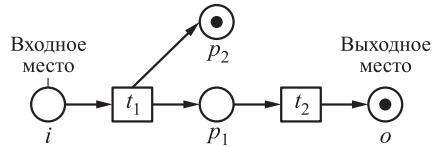


Рис. 3 Размеченная сеть $(N, \{p_2, o\})$

Определение 3 (свойство надежности). Пусть $N = (P, T, F)$ — сеть Петри с входным местом $i \in P$ и выходным местом $o \in P$. Сеть N называется **надежной**, если выполнены следующие условия:

- *корректное завершение*. Для всех достижимых разметок $\forall s \in [N, \{i\}]$, если $o \in s$, то $s = \{o\}$;
- *отсутствие мертвых задач*. Каждый из переходов сети N активируется в какой-либо достижимой разметке $(N, \{i\})$.

Пример на рис. 1 не обладает свойством надежности, так как не выполнено условие корректного завершения.

Перейдем к рассмотрению свойства 2, касающегося запрещенных конструкций (рис. 4). Эти конструкции соответствуют ситуациям, когда происходят подряд выбор и синхронизация. Под *выбором* в контексте сетей Петри понимается конструкция, когда из одного места есть ребра в несколько переходов. Под *синхронизацией* — когда из нескольких мест есть ребра в один и тот же переход. Таким образом, рис. 4, а описывает ситуацию «синхронизация и выбор»: ситуация выбора задачи (место с многими выходными переходами) не должна происходить одновременно с синхронизацией задач (переходы с многими входными местами). Рисунок 4, б описывает ситуацию «выбор и синхронизация»: если происходит синхронизация задач (переход с многими входными местами), то все предшествующие ей задачи должны быть выполнены, т. е. синхронизация задач не должна происходить сразу после операции «ИЛИ».

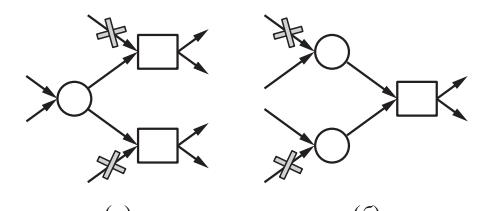


Рис. 4 Запрещенные конструкции согласно свойству 2

зация»: если происходит синхронизация задач (переход с многими входными местами), то все предшествующие ей задачи должны быть выполнены, т. е. синхронизация задач не должна происходить сразу после операции «ИЛИ».

Определение 4 (структурированная сеть Петри). $N = (P, T, F)$ — *структурированная сеть Петри*, если $\forall p \in P, \forall t \in T$, где $(p, t) \in F$, выполнено:

- если $|(p \cdot)| > 1$, то $|(\cdot t)| = 1$;
- если $|(\cdot t)| > 1$, то $|(\cdot p)| = 1$;
- в сети нет неявных мест. Место $p \in P$ называется *неявным*, если для всех достижимых разметок $s' \in [N, s]$ и переходов $t \in (p \cdot)$ выполнено: если $s' \subseteq (\cdot t) \setminus \{p\}$, то $s' \subseteq (\cdot t)$.

Для описания свойства З необходимо ввести дополнительные определения отношений порядка между задачами в логе и срабатывающей последовательности в сети Петри.

Пусть T — конечное множество задач, тогда $\sigma \in T^*$ — *трасса*, $W \subseteq T^*$ — *лог*.

Определение 5 (отношение порядка между задачами в логе). Пусть W — лог над T , и $a, b \in T$. Тогда:

- $a >_W b$ (a предшествует b), если $\exists \sigma = t_1, \dots, t_{n-1}$ и $i \in \{1, \dots, n-2\}$ такие, что $\sigma \in W$ и $t_i = a$, $t_{i+1} = b$;
- $a \rightarrow_W b$ (прямое каузальное отношение), если $a >_W b$ и $b \not>_W a$;
- $a \#_W b$ (задачи не встречаются вместе), если $a \not>_W b$ и $b \not>_W a$;
- $a \parallel_W b$ (потенциальный параллелизм), если $a >_W b$ и $b >_W a$.

Определение 6 (срабатывающая последовательность). Пусть дана размеченная сеть $(N = (P, T, F), s_0)$. Последовательность переходов $\sigma \in T^*$ называется *срабатывающей* для (N, s_0) , если $\exists n \in \mathbb{N}$ такое, что существуют разметки s_1, \dots, s_n и переходы $t_1, \dots, t_n \in T$ такие, что $\sigma = t_1, \dots, t_n$ и для всех i , $0 \leq i < n$ переход t_{i+1} активирован в размеченной сети (N, s_i) и $s_{i+1} = s_i \cup \{(t_{i+1}) \setminus (\cdot t_{i+1})\}$. Обозначение: $(N, s_0)[\sigma]$ Срабатывание последовательности σ ведет к разметке s_n , обозначение: $(N, s_0)[\sigma](N, s_n)$.

Так, в примере на рис. 1 срабатывающей будет последовательность $\sigma = \{t_1, t_2\}$ и после срабатывания разметка $\{i\}$ перейдет в разметку $\{p_2, o\}$, $(N, \{i\})[\delta](N, \{p_2, o\})$.

Определим свойство З.

Определение 7 (полнота лога рабочего процесса). Пусть $N = (P, T, F)$ — надежная сеть Петри. W — лог рабочего процесса для сети N , если $W \subseteq T^*$ и $\forall \sigma \in W$ — срабатывающая последовательность в N , начиная с разметки $\{i\}$ и заканчивая разметкой $\{o\}$. W — полный лог рабочего процесса для сети N , если:

- W — лог рабочего процесса;
- для любого другого лога рабочего процесса W' сети N выполнено: $>'_W \subseteq >_W$;
- $\forall t \in T \exists \sigma \in W : t \in \sigma$ — все переходы покрываются некоторой срабатывающей последовательностью.

Свойство 4 формулируется теоремой в работе [6]. Из теоремы 1 следует, что если существует каузальное отношение между задачами согласно логу рабочего процесса, то в соответствующей сети Петри существует место, соединяющее переходы, соответствующие этим двум задачам.

Теорема 1 [6]. Пусть $N = (P, T, F)$ — надежная сеть Петри и W — полный лог рабочего процесса для сети N . Для всех $a, b \in T$ таких, что $a \rightarrow_W b$, следует

$$(a \cdot) \cap (\cdot b) \neq \emptyset.$$

3 Описание рабочего процесса

Рассматривается задача построения модели рабочего процесса, который представлен логом W . Если такая модель рабочего процесса будет построена, то для нее будет применим алгоритм переобнаружения процесса, предложенный в работе [6]. В данной работе приведен пример рабочего процесса, состоящего из четырех задач, для которого в терминах предложенного в работе [6] математического аппарата не удается построить корректную модель, если под корректностью понимается наличие хотя бы четырех свойств, описанных в разд. 2.

Рассмотрим процесс, состоящий:

- из задачи A , которая всегда выполняется в начале;
- из блока «ИЛИ»: либо задача A , либо задача B , либо никакой задачи;
- из задач C и D , которые всегда выполняются в конце.

Рассмотрим полный лог для такого процесса: $W = \{ACD, AACD, ABCD\}$. Справедливые отношения для лога W :

- отношение предшествия между действиями (задачами, переходами в сети Петри): $A >_W A, A >_W B, A >_W C, B >_W C$ и $C >_W D$;
- прямое каузальное отношение: $A \rightarrow_W B, A \rightarrow_W C, B \rightarrow_W C$ и $C \rightarrow_W D$;
- отношение параллелизма: $A \parallel_W A$;
- задачи не встречаются вместе (симметричное отношение, для краткости не приведены записи вида bRa , если есть запись aRb): $A \#_W D, B \#_W B, B \#_W D, C \#_W C$ и $D \#_W D$.

4 Модели рабочего процесса

Рассмотрим возможные сети Петри для лога W с целью их исследования на наличие четырех свойств, введенных в разд. 2. Можно рассматривать сети Петри, которые представлены в данном разделе, как потенциальные модели процесса, так как трассы лога W могут быть порождены этими сетями. Или, другими словами, сети N_1, N_3, N_4 и N_5 моделируют лог W . Сеть N_2 представлена

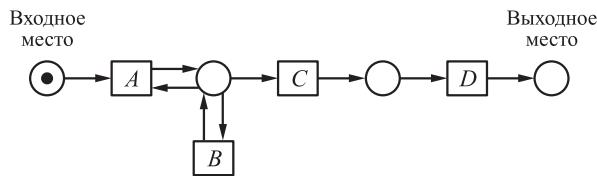


Рис. 5 Сеть N_1 для лога W

как пример неверной работы сложных блоков OR-split и OR-join в классе сетей Петри, предлагаемых авторами [6].

Рассмотрим сеть Петри N_1 (рис. 5) на наличие свойств, определенных в разд. 2.

Утверждение 1. Сеть N_1 надежна.

Доказательство. Проверка свойств по определению надежности.

1. *Корректное завершение.* Так как в сети N_1 нет переходов, увеличивающих число меток (переход с несколькими выходными местами), то если в разметке $\{i\}$ была 1 метка и эта метка дойдет до выходного места, то она будет единственной.
2. *Отсутствие мертвых переходов.* Трасса $ABCD$ активирует каждый из переходов сети N_1 . Построенная сеть N_1 не является структурированной сетью (свойство 2), так как содержит обе запрещенные конструкции (см. рис. 4). На рис. 6 и 7 ребра сети, не участвующие в демонстрации нарушения, изображены пунктирными стрелками, а ребра, демонстрирующие нарушение условия, — толстыми:

 - нарушение условия $\forall(p, t) \in F$, если $|(\cdot p)| > 1$, то $|(\cdot t)| = 1$ (см. рис. 6);
 - нарушение условия $\forall(p, t) \in F$, если $|(\cdot t)| > 1$, то $|(\cdot p)| = 1$ (см. рис. 7).

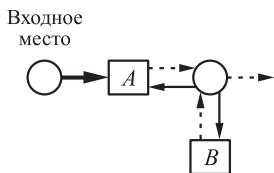


Рис. 6 Синхронизация и выбор в сети N_1

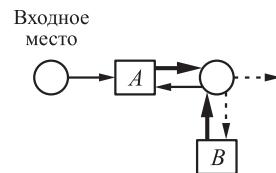


Рис. 7 Выбор и синхронизация в сети N_1

Легко видеть, что сеть моделирует большее число трасс, чем представлено в логе W . Например, срабатывающими последовательностями будут трассы $A \dots ACD$, $AB \dots BCD$, которых нет в логе W . Таким образом, сеть не обладает свойством 3 (полноты лога).

У сети N_1 налицо свойство 4, так как для каждой пары задач в логе W , между которыми есть каузальное отношение, между соответствующими переходами существует место.

Авторы работы [6] полагают, что структурированные сети поддерживают сложные блоки вида OR-split, OR-join, AND-split и AND-join. Так как в логе W есть конструкция «ИЛИ», построим сеть Петри N_2 с переходами OR-split и OR-join (рис. 8).

В рассматриваемом классе сетей Петри условия «И» и «ИЛИ» моделируются с помощью следующих конструкций [7] (рис. 9).

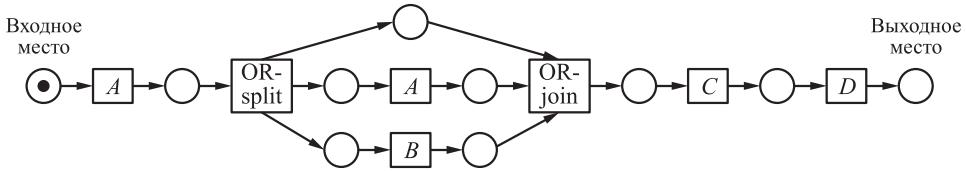


Рис. 8 Сеть N_2 для лога W

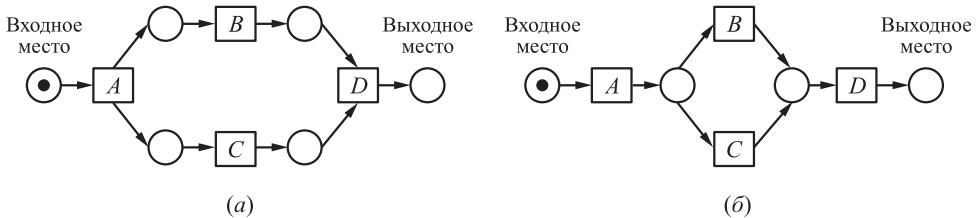


Рис. 9 Конструкции параллельного исполнения (условие «И») (а) и «гэс condition» (условие «ИЛИ») (б)

Таким образом, без наложения дополнительных условий на переходы OR-split и OR-join в сети N_2 эти переходы будут моделировать условие «И», а не условие «ИЛИ», т. е. сеть N_2 будет моделировать не лог W , а лог $W' = \{AABCD, ABACD\}$. Решение данной проблемы может заключаться в рассмотрении более сложного класса сетей Петри с наложением условий на переходы, например класс ингибиторных сетей Петри.

Рассмотрим полный лог W' для сети N_2 . В сети N_2 есть сложные переходы OR-split и OR-join (которые работают в точности, как AND-split и AND-join), и такая сеть все равно не будет корректной, так как не обладает свойством 4:

- лог W' не подразумевает наличия дополнительных переходов, кроме указанных в логике процесса задач (в частности, задач OR-split и OR-join), откуда не может быть определено более полное каузальное отношение \rightarrow (чтобы появились конструкции вида $A \rightarrow \text{OR-join}$);
- в логе W' справедливо каузальное отношение $B \rightarrow_{W'} C$, но при этом в сети N_2 , в срабатывающих разметках, метка может находиться либо между переходами B и OR-join (рис. 10), в таком случае $(\cdot C) = \emptyset$, откуда $(B \cdot) \cap (\cdot C) = \emptyset$, либо метка может находиться между переходами OR-join и C (рис. 11), в таком случае $(B \cdot) = \emptyset$, откуда $(B \cdot) \cap (\cdot C) = \emptyset$.

Используем конструкцию «ИЛИ» с рис. 9, б и построим новую сеть Петри N_3 без переходов OR-split и OR-join (рис. 12). Основной проблемой остается моделирование пустого перехода в условии «ИЛИ», так как места в сетях Петри не могут быть соединены с местами.

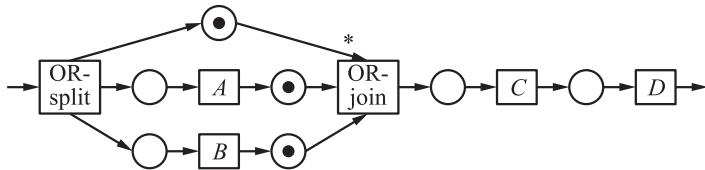


Рис. 10 Часть сети N_2 после срабатывания последовательности AAB или ABA

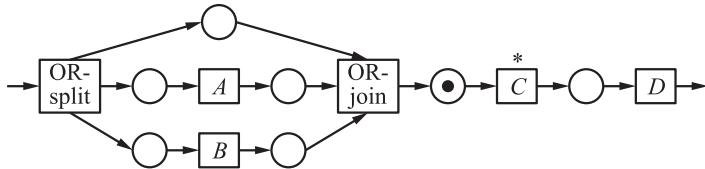


Рис. 11 Часть сети N_2 после срабатывания последовательности $AAABC$ или $ABAC$

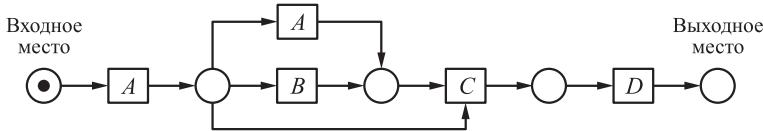


Рис. 12 Сеть N_3 для лога W

Утверждение 2. Сеть N_3 надежна.

Доказательство. Проверка свойств по определению надежности.

1. *Корректное завершение.* В сети N_3 нет переходов, у которых есть несколько выходных мест, поэтому число меток не может увеличиться. Если в сеть Петри не может поступить новая метка, пока находящаяся в ней метка не дойдет до выходного места, то если эта метка окажется в выходном месте, она будет единственной.
2. *Отсутствие мертвых переходов.* Трасса ACD активирует каждый из переходов сети N_3 .

Построенная сеть N_3 не является структурированной сетью (свойство 2), так как содержит обе запрещенные конструкции из рис. 4. На рис. 13 и 14 ребра сети, не участвующие в демонстрации нарушения, изображены пунктирными стрелками, а ребра, демонстрирующие нарушение условия, — жирными стрелками:

- нарушение условия $\forall(p, t) \in F$, если $|(\cdot p)| > 1$, то $|(\cdot t)| = 1$ (см. рис. 13);
- нарушение условия $\forall(p, t) \in F$, если $|(\cdot t)| > 1$, то $|(\cdot p)| = 1$ (см. рис. 14).

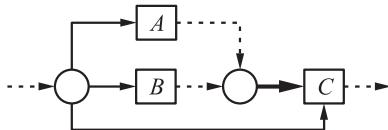


Рис. 13 Синхронизация и выбор в сети N_3

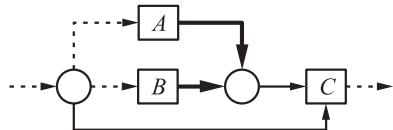


Рис. 14 Выбор и синхронизация в сети N_3

W является полным логом рабочего процесса для сети N_3 (наличие свойства 3), так как содержит все срабатывавшие последовательности для этой сети. Различие в срабатывавших последовательностях для сети N_3 моделируется условием «ИЛИ»: префиксы трасс AAC , ABC или AC . Число различных трасс равно 3, и эти трассы в точности совпадают с логом W .

Лог W и сеть N_3 не обладают свойством 4, касающимся наличия общего места в сети, если в логе между соответствующими задачами справедливо каузальное отношение, и можно доказать более общую теорему.

Теорема 2. *Теорема 1 не выполняется, когда в полном логе W для сети N содержатся повторы задач.*

Доказательство. В доказательство теоремы рассмотрим построенный пример с сетью Петри N_3 и полным логом рабочего процесса W .

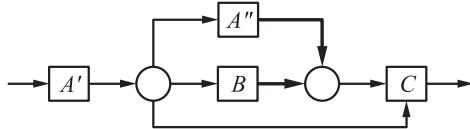


Рис. 15 Часть сети N_3 для иллюстрации противоречия теореме 1 (одинаковые переходы с задачей A размечены дополнительно)

Для лога W справедливо отношение $A \rightarrow_W B$. Согласно теореме 1 из этого отношения должно следовать, что существует общее место для переходов с задачами A и B в сети Петри N_3 . Рассмотрим часть сети N_3 (рис. 15). Можно видеть, что между переходами A'' и B нет общего места.

Проблемы с сетью N_3 связаны с тем, что происходит попытка моделирования условия «ИЛИ», где одна из возможных веток не содержит ни одной задачи; таким образом, выбор и синхронизация должны происходить подряд. Поменяем конструирование пустого перехода в условии «ИЛИ» и построим новую сеть N_4 (рис. 16). Таким образом, у сети N_4 будет заведомо присутство-

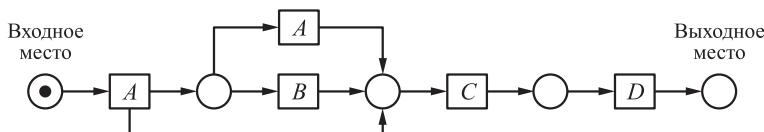


Рис. 16 Сеть N_4 для лога W

вать свойство 2, касающееся того, что сеть не должна содержать запрещенных конструкций. Свойство 3, связанное с полнотой лога W для N_4 , также заведомо верно, по построению.

Утверждение 3. Сеть N_4 некорректно завершается.

Доказательство. Для иллюстрации противоречия рассмотрим срабатывавшую последовательность переходов $\sigma = ACD$, переводящую метку со входного места на выходное место — $(N_4, \{i\})[\sigma](N_4, \{o\})$:

- разметка $\{i\}$ активирует переход A (рис. 17);
- срабатывает переход A , появляются 2 метки, которые активируют 3 перехода (рис. 18);
- срабатывает переход C , меняет местоположение одна метка из двух, активированы 3 перехода (рис. 19);
- срабатывает переход D , метка перемещается на выходное место, в сети еще остается одна метка между переходами A , A и B (рис. 20).

Таким образом, обнаружена достижимая разметка из разметки $\{i\}$ такая, что в этой разметке есть метка, находящаяся на выходном месте, и она не единственная.

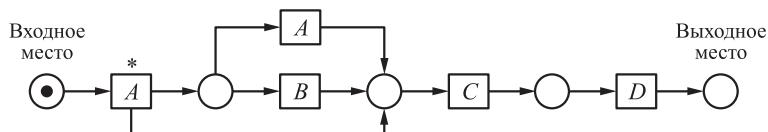


Рис. 17 Часть сети N_4 в начальной разметке. Активированный переход отмечен знаком «*»

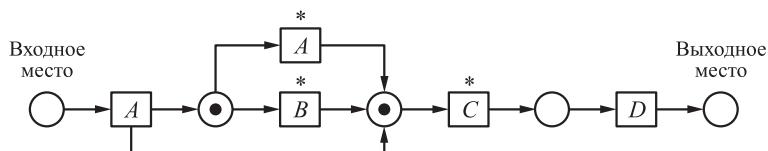


Рис. 18 Часть сети N_4 после срабатывания перехода A

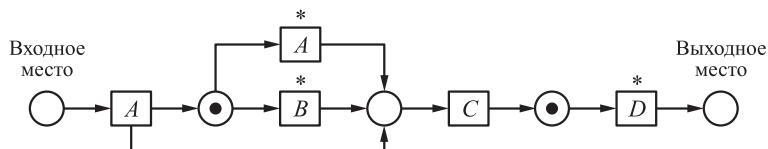


Рис. 19 Часть сети N_4 после срабатывания перехода C

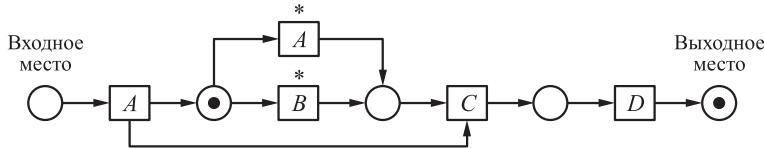


Рис. 20 Часть сети N_4 после срабатывания перехода D

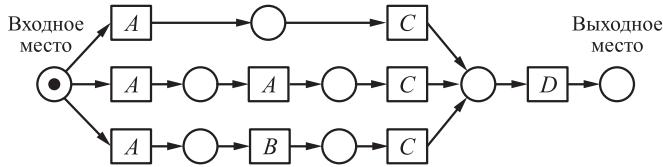


Рис. 21 Сеть N_5 для лога W

Так как сеть N_4 не завершается корректно, она не может быть надежной. Свойство 1 отсутствует. Свойство 4 также отсутствует, так как для сети N_4 справедлива теорема 2.

Увеличим число переходов в сети, для того чтобы получилась каноническая конструкция условия «ИЛИ», как на рис. 9, б. Сеть N_5 представлена на рис. 21.

Сеть N_5 будет обладать свойствами полноты лога и отсутствия запрещенных конструкций по построению. В то же время будет отсутствовать свойство, обеспечивающее выполнение теоремы 1, так как, например, не между каждой парой переходов с задачами A и B существует общее место.

Утверждение 4. Сеть N_5 надежна.

Доказательство. Проверка свойств по определению надежности.

1. *Корректное завершение.* В сети N_5 есть только тип переходов, который не увеличивает число меток. Если в сеть Петри не может поступить новая метка, пока находящаяся в ней метка не дойдет до выходного места, то если эта метка окажется на выходном месте, она будет единственной.
2. *Отсутствие мертвых переходов.* Трассы ACD , $AACD$ и $ABCD$ активируют каждый из переходов сети N_5 .

5 Заключение

В статье проведено исследование предложенного в работе [6] класса сетей Петри для моделирования рабочего процесса. Показано, что при отсутствии ограничения на уникальность задач в логе и наличии в модели переходов, которые не соответствуют никакой задаче в логе, построение модели, которая бы обладала хотя бы рассматриваемыми четырьмя свойствами, является сложной

Сравнение сетей Петри для одного процесса

Сеть	Надежность	Нет запрещенных конструкций	Полнота лога	Сохранение отношения каузальности
N_1	+	—	—	+
N_2	—*	—*	—*	—*
N_3	+	—	+	—
N_4	—	+	+	—
N_5	+	+	+	—

задачей, где выполнение одних требований часто противоречит выполнению других. Сравнительные характеристики рассмотренных сетей по свойствам представлены в таблице.

Также в работе показано, что переходы OR-split и OR-join без наложения дополнительных условий на рассматриваемый класс сетей Петри будут работать в точности, как переходы AND-split и AND-join, и будут моделировать другой процесс. Поэтому в таблице результаты сравнений по требованиям для сети N_2 помечены знаком «*».

Таким образом, можно говорить о том, что алгоритм обнаружения процесса по логу из статьи [6] справедлив для еще более узкого класса рабочих процессов, чем предполагается в [6]. Более точно, в логе рабочего процесса задачи должны быть уникальными, не должно быть переходов по пустому условию, не должно быть моделирования условий «ИЛИ».

Литература

1. Agrawal R., Gunopulos D., Leymann F. Mining process models from workflow logs // Advances in database technology / Eds. H. J. Schek, G. Alonso, F. Saltor, I. Ramos. — Lecture notes in computer science ser. — Springer, 1998. Vol. 1337. P. 467–483.
2. Cook J. E., Wolf A. L. Discovering models of software processes from event-based data // ACM T. Softw. Eng. Meth., 1998. Vol. 7. No. 3. P. 215–249.
3. Mannila H., Rusakov D. Decomposition of event sequences into independent components // SIAM Conference (International) on Data Mining Proceedings. — Society for Industrial and Applied Mathematics, 2001. P. 1–17.
4. Schimm G. Process miner — a tool for mining process schemes from event-based data // Logics in artificial intelligence / Eds. S. Flesca, S. Greco, G. Ianni, N. Leone. — Lecture notes in computer science ser. — Springer, 2002. Vol. 2424. P. 525–528.
5. Herbst J. Ein induktiver Ansatz zur Akquisition und Adaption von Workflow-Modellen. — Berlin: Tenea Verlag Ltd., 2004. 284 p.
6. Van der Aalst W., Weijters T., Maruster L. Workflow mining: Discovering process models from event logs // IEEE T. Knowl. Data En., 2004. Vol. 16. No. 9. P. 1128–1142.
7. Lectures on Petri nets I: Basic models: Advances in Petri nets / Eds. W. Reisig, G. Rozenberg. — Berlin–Heidelberg: Springer-Verlag, 1998. 691 p.

8. Питерсон Д. Теория сетей Петри и моделирование систем / Пер. с англ. — М.: Мир, 1984. 264 с. (Peterson J. L. Petri net theory and the modeling of systems. — Prentice-Hall, 1981. 290 p.)

Поступила в редакцию 11.09.20

CONSTRUCTING PROCESS MODELS REPRESENTED BY SIMPLE PETRI NETS

I. Yu. Teryokhina¹, A. A. Grusho², E. E. Timonina², and S. Ya. Shorgin²

¹Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation

²Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper deals with the problem of “workflow mining.” Workflow mining is numerous techniques for discovering processes’ models represented by their workflow log. The paper considers process models in terms of simple Petri nets. It is shown that constructing a correct model when a process contains equal tasks is not always an attainable goal. Moreover, it was revealed that in the case when a model has transitions with no correspondence to any process task, the relation between the causal relations detected in the log and the presence of places connecting transitions in the Petri net is violated.

Keywords: Petri nets; workflow mining; process modeling

DOI: 10.14357/08696527200406

Acknowledgments

The paper was partially supported by RFBR, project No. 18-07-00274.

References

1. Agrawal, R., D. Gunopulos, and F. Leymann. 1998. Mining process models from workflow logs. *Advances in database technology*. Eds. H. J. Schek, G. Alonso, F. Saltor, and I. Ramos. Lecture notes in computer science ser. Springer. 1337:467–483.
2. Cook, J. E., and A. L. Wolf. 1998. Discovering models of software processes from event-based data. *ACM T. Softw. Eng. Meth.* 7(3):215–249.
3. Mannila, H., and D. Rusakov. 2001. Decomposition of event sequences into independent components. *SIAM Conference (International) on Data Mining Proceedings*. Society for Industrial and Applied Mathematics. 1–17.

4. Schimm, G. 2002. Process miner — a tool for mining process schemes from event-based data. *Logics in artificial intelligence*. Eds. S. Flesca, S. Greco, G. Ianni, and N. Leone. Lecture notes in computer science ser. Springer. 2424:525–528.
5. Herbst, J. 2004. *Ein induktiver ansatz zur akquisition und adaption von workflow-modellen*. Berlin: Tenea Verlag Ltd. 284 p.
6. Van der Aalst, W., T. Weijters, and L. Maruster. 2004. Workflow mining: Discovering process models from event logs. *IEEE T. Knowl. Data En.* 16(9):1128–1142.
7. Reisig, W., and G. Rozenberg, eds. 1998. *Lectures on Petri nets I: Basic models: Advances in Petri nets*. Berlin–Heidelberg: Springer-Verlag. 691 p.
8. Peterson, J. L. 1981. *Petri net theory and the modeling of systems*. Prentice-Hall. 290 p.

Received September 11, 2020

Contributors

Teryokhina Irina Yu. (b. 1994)— PhD student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskie Gory, GSP-1, Moscow 119991, Russian Federation; irina_teryokhina@mail.ru

Grusho Alexander A. (b. 1946)— Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; grusho@yandex.ru

Timonina Elena E. (b. 1952)— Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; eltimon@yandex.ru

Shorgin Sergey Ya. (b. 1952)— Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; sshorgin@ipiran.ru

КОМПЛЕКСНОЕ ОБЕСПЕЧЕНИЕ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ ЧАСТНЫХ ОБЛАЧНЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕД^{*}

*А. А. Грушо¹, А. В. Николаев², В. О. Писковский³, В. В. Сенчило⁴,
Е. Е. Тимонина⁵*

Аннотация: Рассматривается задача построения безопасного корпоративного облака, в котором используются небезопасные компоненты. Наиболее важным представляется решение для удаленной работы сотрудников организации. Решение основано на выделении простого подоблака, реализующего безопасные взаимодействия удаленных сотрудников между собой, с облачными сервисами и ресурсами глобальной сети Интернет.

Keywords: информационная безопасность; облачные вычисления; безопасность удаленной работы на небезопасных терминалах

DOI: 10.14357/08696527200407

1 Введение

Различные организации все шире используют облачные вычисления. Вынос точки доступа в частное облако за пределы контролируемой зоны (КЗ) всегда связан с целым рядом рисков информационной безопасности (ИБ). В качестве такой точки доступа может быть использован ноутбук, персональный компьютер, планшет или телефон. Удаленное рабочее место может содержать деловую переписку, важные документы, статистику. Его легко потерять, лишиться при транспортировке и пр. Таким образом, обработка важной или секретной информации на таком рабочем месте вне КЗ проблематична. Кроме несанкционированного доступа к конфиденциальной информации, хранимой на рабочей станции, подключение ее к домашней сети может завершиться взломом и инфицированием с последующим неконтролируемым проникновением в офисную информационную среду или инфраструктуру организации. Без профессиональных средств

^{*}Работа частично поддержанна РФФИ (проект 18-29-03081).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, grusho@yandex.ru

²Федеральный исследовательский центр химической физики им. Н. Н. Семёнова Российской академии наук, gentoorion@mail.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vrupr80@yandex.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, volodias@mail.ru

⁵Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, eltimon@yandex.ru

защиты обнаружить подобную атаку практически невозможно, как и невозможно предотвратить использование сторонних средств хранения и обмена информации, интернет-ресурсов. Служба ИБ компании обычно не имеет возможности про-контролировать удаленные рабочие места, что делает такие сценарии реализации угроз еще более критичными с точки зрения последствий. Контроль действий при обмене информацией становится еще более сложной задачей. Очевидно, что для обеспечения удаленного рабочего места, безопасного с точки зрения интересов компании, нужен новый подход к организации эффективной системы защиты, которая создавалась бы с учетом вышеперечисленных угроз. При этом система должна быть простой в эксплуатации, а лучше — скрытой, управление не должно требовать высокой квалификации обслуживающего персонала.

2 Архитектура защиты

В течение ряда лет в Российской Федерации проводились исследования по разработке архитектуры защищенной информационной системы (ИС) из небезопасных компонентов [1, 2]. Ряд этих результатов можно применить и к построению корпоративного облака. В работе предложено в отдельное подоблако вынести:

- удаленное взаимодействие сотрудников корпорации между собой;
- удаленное использование ресурсов компании;
- взаимодействие с ресурсами глобальной сети.

Такое подоблако позволит локализовать и изолировать опасные функции, связанные с взаимодействиями и сетевыми обменами. Отметим, что остальная часть облачных сервисов также окажется в изоляции, что позволит использовать в облаке небезопасные компоненты. Идея максимальной изоляции подсистем, построенных из небезопасных элементов, стала одной из основных в теории построения защищенных систем из незащищенных элементов [3, 4].

Этот подход частично пересекается с новой разработкой американских учебных, а именно: с концепцией в области кибербезопасности — «нулевое доверие» (Zero Trust, или ZT) [5]. Этот подход предполагает, что каждая сеть взломана, каждая машина скомпрометирована и каждый пользователь (вольно или невольно) становится либо нарушителем, либо объектом интереса нарушителя. Никому и ничему в сети с нулевым доверием нельзя доверять, пока не доказано обратное.

Национальный институт стандартов и технологий США (NIST) [5] определяет ZT как парадигму кибербезопасности, направленную на защиту ресурсов в предположении, что уровень доверия никогда не предоставляется по умолчанию, а постоянно оценивается. Архитектура сети, соответствующая ZT, представляет собой план кибербезопасности предприятия, который использует концепции нулевого доверия и охватывает отношения компонентов, планирование рабочего

процесса и политики доступа. Таким образом, ZT-безопасность — это совокупность взаимосвязанных политик, практик, программного обеспечения (ПО) и оборудования.

ZT не выступает альтернативой традиционным методам безопасности, а является развитием хорошо известных и зарекомендовавших себя технологий. Повысить уровень доверия к среде, в которой работает пользователь, ZT предлагает с помощью ряда известных мер, в числе которых:

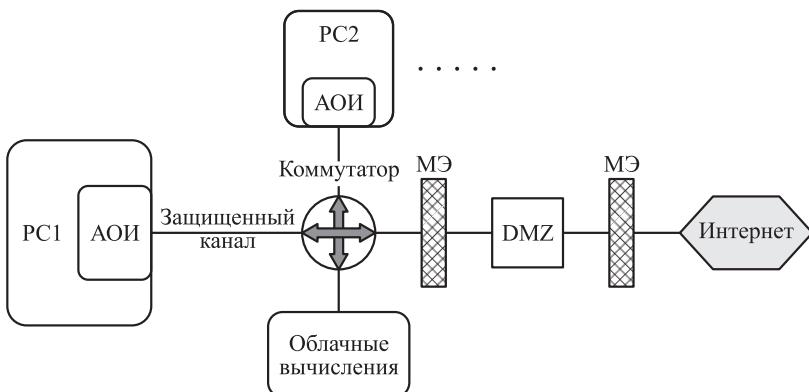
- установка и поддержка антивирусного ПО;
- установка и корректная настройка межсетевого экрана;
- корректная настройка процессов обновления системного и прикладного ПО;
- установка аппаратного модуля доверенной загрузки.

Для создания сети в рассматриваемых предположениях необходимо наличие ряда автоматизированных служб и компонентов поддержки безопасности. Все эти элементы служат инструментами или процедурами, которые уже успешно существуют, эксплуатируются и повторяют концепцию программно-коммутируемой сети (ПКС). Так, компонент принятия решения о политике (PDP — Policy Decision Point) решает, безопасен ли компьютер или веб-трафик, и после этого соответственно предоставляет или отменяет доступ. Механизм политик безопасности (ПБ) [6] использует все доступные, в том числе и внешние, источники данных, применимые к потребностям организации для определения безопасности на основе ПБ. Так, решение о безопасности сетевого компонента принимается на основе информации, предоставляемой системами. Система непрерывной диагностики и компенсации последствий считается главной. Система собирает информацию о текущем состоянии сетевых компонентов и применяет обновления к конфигурации и программным компонентам, предоставляет информацию об элементе, отправляющем запрос на доступ, например работает ли он с соответствующей исправленной операционной системой и приложениями или есть ли у актива какие-либо известные уязвимости.

Основным недостатком описанного подхода является наличие сложной интеллектуальной системы отслеживания и динамического принятия решений. В рамках достижений информационных технологий в РФ такую систему создать сложно. Второй недостаток американского подхода состоит в том, что он опирается на стандартный набор механизмов безопасности, которые не всегда дают надежность защиты. В отечественных разработках предложена концепция контроля соединений распределенной системы на основе метаданных, которая позволяет обеспечить более высокий уровень ИБ [7].

3 Безопасные взаимодействия

В качестве примера применения частного подоблака для организации безопасных взаимодействий рассмотрим реализацию частного облака, представленного на рисунке.



Архитектура защищенного корпоративного облака с помощью безопасного коммуникационного подоблака

В качестве пользовательского терминала используется виртуальная машина (или контейнер), на которой реализован клиент для терминальных служб, развернутых в частном облаке. Привлекает идея загрузки в недоверенный терминал пользователя «чистой» рабочей среды с внешнего носителя или облачного ресурса. Будем называть устройство безопасного доступа пользовательского терминала к облачным ресурсам *абонентским облачным интерфейсом* (АОИ). Рабочая станция (РС) пользователя не имеет возможности непосредственного выхода в глобальную сеть. Единственным каналом служит защищенный канал из АОИ, соединяющий РС1 с частным подоблаком и реализующий тотальное шифрование информации, передаваемой по этому каналу на уникальном для АОИ ключе.

В соответствии с вышесказанным, РС1 с установленным АОИ запрашивает нужное взаимодействие в подоблаке контроля взаимодействий. Между коммутатором, служащим точкой доступа в облако, и АОИ устанавливается защищенное соединение на ключе, используемом только между этой РС и шифратором на коммутаторе. Каждый легальный АОИ имеет уникальный симметричный ключ для связи с коммутатором. Успешное расшифрование обращения к коммутатору служит надежной аутентификацией АОИ и инициирует продолжение организации взаимодействия.

Если взаимодействие предполагает связь с другим сотрудником корпорации, то осуществляется его вызов через аналогичный терминал этого сотрудника на его симметричном ключе. Для их взаимодействия потребуется перешифрование обмена информацией с одного ключа на другой. Описанная схема взаимодействия — простейшая, она оставляет возможность контроля взаимодействия. Для маршрутизации трафика внутри облака коммутаторы строят индивидуальные для каждого узла защищенные соединения с помощью развернутой инфраструктуры средств шифрования предприятия.

4 Заключение

Построение современных корпоративных систем с защищенными терминалами удаленных друг от друга пользователей может опираться на незащищенные РС и мобильные устройства, в которые встроен единственный доверенный блок, реализующий интерфейс небезопасного компьютера без возможности выхода в сеть и сетевого интерфейса с единственным адресом взаимодействия. Этот адрес предполагает связь с примитивным подоблаком корпоративной системы, реализующей коммутатор для организации взаимодействия удаленных пользователей друг с другом, с сервисами корпоративного облака и ресурсами внешней глобальной сети. В зависимости от ценностей активов организаций может использоваться гарантированное или легкое шифрование.

Литература

1. *Grusho A. A. Composition of the trusted computer security using untrusted elements // Probabilistic methods in discrete mathematics / Eds. V. F. Kolchin, V. Ya. Kozlov, Yu. L. Pavlov, Yu. V. Prokhorov. — Utrecht, Netherlands: VSP, 1997. P. 231–236.* doi: 10.1515/9783112314074-022.
2. *Тимонина Е. Е. Анализ угроз скрытых каналов и методы построения гарантированно защищенных распределенных автоматизированных систем: Дис. . . . докт. техн. наук. — М., 2004. 204 с.*
3. *Грушо А. А., Володин А. В., Тимонина Е. Е. Безопасный интерфейс с глобальной сетью из ненадежных в смысле безопасности элементов // Информационные технологии в науке, образовании, телекоммуникации, бизнесе (весенняя сессия): Тр. Междунар. конф. — Запорожье, Украина: Лаборатория издательских технологий и компьютерной графики Запорожского государственного университета, 2001. С. 149–151.*
4. *Грушо А. А., Тимонина Е. Е. Гарантированно защищенные базы данных, построенные на недоверенных с точки зрения безопасности элементах // Проблемы безопасности и противодействия терроризму: Мат-лы конф. — М.: МЦНМО, 2006. С. 335–348.*
5. *Borchert O., Mitchell S., Connelly S., Rose S. Zero trust architecture // Draft (2nd) NIST Special Publication 800-207, 2000. <https://csrc.nist.gov/publications/detail/sp/800-207/final>.*
6. *Грушо А. А., Применко Э. А., Тимонина Е. Е. Теоретические основы компьютерной безопасности. — М.: Академия, 2009. 272 с.*
7. *Grusho A. A., Timonina E. E., Shorgin S. Ya. Modelling for ensuring information security of the distributed information systems // 31th European Conference on Modelling and Simulation Proceedings. — Dudweiler, Germany: Digitaldruck Pirrot GmbH, 2017. P. 656–660. <http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstatECMS20170026.pdf>.*

Поступила в редакцию 12.09.20

END-TO-END INFORMATION SECURITY OF PRIVATE CLOUD COMPUTING

**A. A. Grusho¹, A. V. Nikolaev², V. O. Piskovski¹, V. V. Senchilo¹,
and E. E. Timonina¹**

¹Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation

²N. N. Semenov Federal Research Center for Chemical Physics of the Russian Academy of Sciences, 4 Kosygin Str., Moscow 119991, Russian Federation

Abstract: The paper discusses building of a secure enterprise cloud that uses unsecure components. The most important solution is for the organization's employees to work remotely. The solution is based on provision of a simple subcloud that implements secure interactions of remote employees among themselves, cloud services, and interaction with resources in the global Internet.

Keywords: information security; cloud computing; security of remote work on unsecure terminals

DOI: 10.14357/08696527200407

Acknowledgments

The paper was partially supported by RFBR, project No. 18-29-03081.

References

1. Grusho, A. A. 1997. Composition of a trusted computer security system on the base of untrusted elements. *Probabilistic methods in discrete mathematics*. Eds. V. F. Kolchin, V. Ya. Kozlov, Yu. L. Pavlov, and Yu. V. Prokhorov. Utrecht, Netherlands: VSP. 231–236. doi: 10.1515/9783112314074-022.
2. Timonina, E. E. 2004. Analiz ugroz skrytykh kanalov i metody postroeniya garantirovanno zashchishchennykh raspredelennykh avtomatizirovannykh sistem [The analysis of threats of covert channels and methods of creation of guaranteed protected distributed automated systems]. Moscow: Russian State University for the Humanities. D.Sc. Diss. 204 p.
3. Grusho, A. A., A. V. Volodin, and E. E. Timonina. 2001. Bezopasnyy interfeys s global'noy set'yu iz nenadezhnykh v smysle bezopasnosti elementov [Secure interface with WAN from untrusted elements]. *Informatsionnye tekhnologii v naуke, obrazovanii, telekommunikatsii, biznese (vesennaya sessiya): Tr. Mezhdunar. konf.* [Conference (International) “Information Technologies in Science, Education, Telecommunications, Business (spring session)” Proceedings]. Zaporozhye: Laboratory of Publishing Technologies and Computer Graphics of Zaporozhye State University. 149–151.

4. Grusho, A. A., and E. E. Timonina. 2006. Garantirovanno zashchishchennye bazy dannykh, postroennye na nedoverennykh s tochki zreniya bezopasnosti elementakh [Guaranteed secure databases built on insecure elements]. *Problemy bezopasnosti i protivodeystviya terrorizmu: Mat-ly konf.* [Conference “Security and Counter-Terrorism Issues” Proceedings]. Moscow: MCCME. 335–348.
5. Borchert, O., S. Mitchell, S. Connelly, and S. Rose. 2000. SP 800-207: Zero trust architecture. NIST. Available at: <https://csrc.nist.gov/publications/detail/sp/800-207/final> (accessed September 12, 2020).
6. Grusho, A., E. Primenko, and E. Timonina. 2009. *Teoreticheskie osnovy komp'yuternoy bezopasnosti* [Theoretical bases of computer security]. Moscow: Academy. 272 p.
7. Grusho, A. A., E. E. Timonina, and S. Ya. Shorgin. 2017. Modelling for ensuring information security of the distributed information systems. *31th European Conference on Modelling and Simulation Proceedings*. Dudweiler, Germany: Digitaldruck Pirrot GmbH. 656–660. Available at: <http://www.scs-europe.net/dlib/2017/ecms2017acceptedpapers/0656-probstatECMS20170026.pdf> (accessed September 12, 2020).

Received September 12, 2020

Contributors

Grusho Alexander A. (b. 1946) — Doctor of Science in physics and mathematics, professor, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; grusho@yandex.ru

Nikolaev Andrei V. (b. 1973) — Candidate of Science (PhD) in physics and mathematics, scientist, N. N. Semenov Federal Research Center for Chemical Physics of the Russian Academy of Sciences, 4 Kosygin Str., Moscow 119991, Russian Federation; gentoorion@mail.ru

Piskovski Viktor O. (b. 1963) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; vpv80@yandex.ru

Senchilo Vladimir V. (b. 1963) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; volodias@mail.ru

Timonina Elena E. (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences; 44-2 Vavilov Str., Moscow 119133, Russian Federation; eltimon@yandex.ru

ОБ АЛГОРИТМАХ ДЕКОДИРОВАНИЯ ОБОБЩЕННЫХ КОДОВ РИДА–СОЛОМОНА

С. М. Рацееv¹, О. И. Череватенко²

Аннотация: Рассматриваются алгоритмы декодирования обобщенных кодов Рида–Соломона (РС), построенные на основе алгоритмов для кодов РС. Приводится алгоритмы Гао, Сугиямы и Берлекэмпа–Месси (алгоритм Питерсона–Горенстейна–Цирлера). Первый из данных алгоритмов относится к алгоритмам бессиндромного декодирования, остальные — к алгоритмам синдромного декодирования. Актуальность данных алгоритмов состоит в том, что они применимы для декодирования кодов Гоппы, которые лежат в основе некоторых перспективных постквантовых криптосистем. При этом данные алгоритмы применимы для кодов Гоппы над произвольным полем в отличие от хорошо известного алгоритма декодирования Паттерсона для двоичных кодов Гоппы.

Ключевые слова: помехоустойчивые коды; коды Рида–Соломона; коды Гоппы; декодирование кода

DOI: 10.14357/08696527200408

1 Введение

Пусть $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{n-1})$, где α_i – различные элементы поля $\text{GF}(q)$, $y = (y_0, y_1, \dots, y_{n-1})$ — ненулевые (необязательно различные) элементы из $\text{GF}(q)$. Тогда обобщенный $[n, k, d = n - k + 1]$ -код РС, обозначаемый $\text{GRS}_k(\alpha, y)$, состоит из всех кодовых векторов вида

$$u = (y_0 b(\alpha_0), y_1 b(\alpha_1), \dots, y_{n-1} b(\alpha_{n-1})) ,$$

где $b(x)$ — информационные многочлены над полем $\text{GF}(q)$ степени не выше $k - 1$. Если $n = q - 1$, вектор y состоит из единиц и $\alpha_i = \alpha^i$, $i = 0, 1, \dots, n - 1$, где α — примитивный элемент поля $\text{GF}(q)$, то в этом случае получаем код РС. Шиозаки [1] и Гао [2] предложили очень простой алгоритм декодирования кодов РС, асимптотическая сложность которого оценивается величиной $O(n(\log n)^2)$ и совпадает со сложностью лучших алгоритмов декодирования кодов РС [3–5]: Питерсона–Горенстейна–Цирлера (сложность алгоритма $O(n^3)$) [6], Берлекэмпа–Месси

¹Ульяновский государственный университет, ratseevsm@mail.ru

²Ульяновский государственный педагогический университет имени И. Н. Ульянова, choi2008@yandex.ru

(сложность алгоритма $O(n^2)$ [6]) и Сугиямы–Касахары–Хирасавы–Намекавы (сложность алгоритма $O(n^2)$ [6]).

Заметим, что, в отличие от кодов РС, в обобщенных кодах РС одна из компонент вектора α может быть нулевой, что нужно учитывать для некоторых алгоритмов декодирования. В данной работе рассматриваются алгоритмы декодирования для обобщенных кодов РС, построенные на основе алгоритмов для кодов РС. Актуальность таких алгоритмов декодирования состоит в том, что все алгоритмы декодирования для обобщенных кодов РС можно применять для декодирования кодов Гоппы, при этом именно на основе кодов Гоппы строятся некоторые перспективные постквантовые криптосистемы [7].

2 Декодирование обобщенных кодов Рида–Соломона на основе алгоритма Гао

При описании следующего алгоритма будем следовать работе [8]. Пусть обобщенный код РС $\text{GRS}_k(\alpha, y)$ длины n задается векторами α и y . Определим многочлен $m(x) = (x - \alpha_0)(x - \alpha_1) \cdots (x - \alpha_{n-1})$. Пусть кодовый вектор u получен с помощью кодирования информационного многочлена $b(x) = b_0 + b_1x + \cdots + b_{k-1}x^{k-1}$:

$$u = (y_0b(\alpha_0), y_1b(\alpha_1), \dots, y_{n-1}b(\alpha_{n-1})).$$

Пусть $v = u + e$ — полученный вектор, где e — вектор ошибок, $X_1 = \alpha_{i_1}, \dots, X_t = \alpha_{i_t}$ — локаторы ошибок; $Y_1 = e_{i_1}, \dots, Y_t = e_{i_t}$ — значения ошибок. В данном алгоритме многочлен локаторов ошибок запишем в виде:

$$\sigma(x) = (x - X_1) \cdots (x - X_t).$$

Если ошибок не было, то будем полагать, что $\sigma(x) = 1$.

Если $v_i = u_i$, то $v_i = y_i b(\alpha_i)$. Если $v_i \neq u_i$, то на позиции i произошла ошибка, поэтому $\sigma(\alpha_i) = 0$. Из этого следует, что $\sigma(\alpha_i)y_i^{-1}v_i = \sigma(\alpha_i)b(\alpha_i)$, $i = 0, 1, \dots, n-1$. Обозначим $p(x) = \sigma(x)b(x)$. Тогда

$$\sigma(\alpha_i)y_i^{-1}v_i = p(\alpha_i), \quad i = 0, 1, \dots, n-1.$$

Построим интерполяционный многочлен Лагранжа $f(x)$ степени не выше $n-1$, проходящий через точки $(\alpha_0, y_0^{-1}v_0)$, $(\alpha_1, y_1^{-1}v_1)$, \dots , $(\alpha_{n-1}, y_{n-1}^{-1}v_{n-1})$:

$$f(\alpha_i) = y_i^{-1}v_i, \quad i = 0, 1, \dots, n-1, \quad \deg f(x) \leq n-1.$$

Тогда из равенств $\sigma(\alpha_i)f(\alpha_i) = p(\alpha_i)$, $i = 0, 1, \dots, n-1$, получаем сравнение:

$$\sigma(x)f(x) \equiv p(x) (\text{mod } m(x)).$$

Алгоритм 1 (алгоритм декодирования Гао для обобщенных кодов Рида–Соломона)

Вход: принятый вектор v .

Выход: исходный информационный вектор b , если произошло не более $[(d - 1)/2]$ ошибок.

1. *Интерполяция.* Строится интерполяционный многочлен $f(x)$, для которого $f(\alpha_i) = y_i^{-1}v_i$, $i = 0, 1, \dots, n - 1$.

2. *Незаконченный обобщенный алгоритм Евклида.* Пусть $r_{-1}(x) = m(x)$, $r_0(x) = f(x)$, $v_{-1}(x) = 0$, $v_0(x) = 1$. Производится последовательность действий обобщенного алгоритма Евклида

$$r_{i-2}(x) = r_{i-1}(x)q_{i-1}(x) + r_i(x); \quad v_i(x) = v_{i-2}(x) - v_{i-1}(x)q_{i-1}(x), \quad i \geq 1,$$

до тех пор, пока не будет достигнут такой $r_j(x)$, для которого

$$\deg r_{j-1}(x) \geq \frac{n+k}{2}; \quad \deg r_j(x) < \frac{n+k}{2}.$$

3. *Деление.* Информационный многочлен равен $b(x) = r_j(x)/v_j(x)$.

Теорема 1. Если в кодовом векторе произошло не более $[(d - 1)/2]$ ошибок, то алгоритм декодирования 1 всегда приводит к единственному решению, а именно: к исходному информационному вектору b .

Доказательство данной теоремы аналогично доказательству аналогичной теоремы из работы [8] для кодов РС.

Пример 1. Рассмотрим обобщенный код РС над полем GF(7) с параметрами $n = 7$, $k = 3$, $d = 5$, $\alpha = (0, 1, 2, 3, 4, 5, 6)$, $y = (2, 1, 3, 1, 4, 1, 5)$. Так как вектор α содержит все элементы поля GF(7), то $m(x) = x^7 - x$. Ниже приведены матрица Вандермонда V на основе вектора α , обратная к ней матрица V^{-1} и диагональная матрица Y на основе вектора y :

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 0 & 1 & 4 & 2 & 2 & 4 & 1 \\ 0 & 1 & 1 & 6 & 1 & 6 & 6 \\ 0 & 1 & 2 & 4 & 4 & 2 & 1 \\ 0 & 1 & 4 & 5 & 2 & 3 & 6 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}; \quad V^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 6 \\ 0 & 6 & 6 & 6 & 6 & 6 & 6 \\ 0 & 3 & 5 & 6 & 3 & 5 & 6 \\ 0 & 2 & 3 & 1 & 5 & 4 & 6 \\ 0 & 5 & 3 & 6 & 5 & 3 & 6 \\ 0 & 4 & 5 & 1 & 3 & 2 & 6 \\ 0 & 1 & 6 & 1 & 6 & 1 & 6 \end{pmatrix}; \quad Y = \text{Diag}(2, 1, 3, 1, 4, 1, 5).$$

Заметим, что первые $k = 3$ строки матрицы YV образуют порождающую матрицу G кода GRS₃(α, y).

Пусть $b = (2, 4, 1)$ — информационный вектор, который соответствует многочлену $b(x) = 2 + 4x + x^2$. После кодирования вектора b получаем кодовый вектор

$$u = (y_0b(0), y_1b(1), \dots, y_6b(6)) = (4, 0, 0, 2, 3, 5, 2) = (b_0, b_1, b_2, 0, 0, 0, 0,)VY.$$

Пусть вектор ошибок равен $e = (5, 0, 3, 0, 0, 0, 0)$. Тогда на приемном конце будет получен вектор $v = u + e = (2, 0, 3, 2, 3, 5, 2)$. Применим алгоритм декодирования 1.

1. *Интерполяция.* Вычисляем коэффициенты многочлена $f(x) = f_0 + f_1x + \dots + f_6x^6$:

$$(f_0, f_1, \dots, f_6) = vY^{-1}V^{-1} = (1, 0, 6, 6, 3, 5, 0).$$

2. *Незаконченный обобщенный алгоритм Евклида.* Полагаем $r_{-1}(x) = m(x)$, $r_0(x) = f(x)$, $v_{-1}(x) = 0$, $v_0(x) = 1$. После применения первого шага обобщенного алгоритма Евклида

$$\begin{aligned} r_{-1}(x) &= r_0(x)(x + 3x^2) + 5x + 4x^2 + x^3 + 4x^4; \\ r_1(x) &= 5x + 4x^2 + x^3 + 4x^4; \\ v_1(x) &= -(x + 3x^2) = 6x + 4x^2 \end{aligned}$$

процесс останавливается, так как $\deg r_0(x) = 5$, $\deg r_1(x) = 4$, $(n+k)/2 = 5$.

3. *Деление.* Исходный информационный многочлен равен $b(x) = r_1(x)/v_1(x)$.

3 Декодирование обобщенных кодов Рида-Соломона на основе алгоритма Сугиямы

В следующей теореме приводится вид проверочной матрицы кода $\text{GRS}_k(\alpha, y)$ (см., например, [3]), которая необходима для рассмотренного ниже алгоритма декодирования.

Теорема 2. *Код, дуальный $\text{GRS}_k(\alpha, y)$ -коду, является $\text{GRS}_{n-k}(\alpha, w)$ -кодом для некоторого вектора w , причем в качестве вектора w можно взять $w = (w_0, w_1, \dots, w_{n-1})$, где*

$$w_i = \frac{1}{y_i \prod_{j \neq i} (\alpha_i - \alpha_j)}, \quad i = 0, 1, \dots, n-1.$$

Из теоремы 2 следует, что проверочная матрица H кода $\text{GRS}_k(\alpha, y)$ равна порождающей матрице кода $\text{GRS}_{n-k}(\alpha, w)$:

$$H = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_0 & \alpha_1 & \cdots & \alpha_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_0^{n-k-1} & \alpha_1^{n-k-1} & \cdots & \alpha_{n-1}^{n-k-1} \end{pmatrix} \begin{pmatrix} w_0 & 0 & \cdots & 0 \\ 0 & w_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{n-1} \end{pmatrix}. \quad (1)$$

Матрица H содержит $n - k = d - 1$ строк и n столбцов.

Пусть задан код $\text{GRS}_k(\alpha, y)$ с проверочной матрицей вида (1) с кодовым расстоянием $d \geq 2t + 1$. Так как $r = n - k = d - 1 \geq 2t$, то будем рассматривать синдромный вектор вида $(S_0, S_1, \dots, S_{2t-1})$, равный произведению вектора v на транспонированную матрицу, полученную из первых $2t$ строк матрицы H . Распишем это более подробно.

Пусть после отправления вектора u на приемном конце принят вектор $v = u + e$, где e — вектор ошибок веса не более t . Пусть ошибки произошли на позициях i_1, \dots, i_t . Вычисляя синдромный вектор, получаем:

$$S = vH^T = eH^T = \begin{pmatrix} e_{i_1}w_{i_1} + \dots + e_{i_t}w_{i_t} \\ e_{i_1}w_{i_1}\alpha_{i_1} + \dots + e_{i_t}w_{i_t}\alpha_{i_t} \\ \vdots \\ e_{i_1}w_{i_1}\alpha_{i_1}^{r-1} + \dots + e_{i_t}w_{i_t}\alpha_{i_t}^{r-1} \end{pmatrix}^T.$$

Пусть $X_1 = \alpha_{i_1}, X_2 = \alpha_{i_2}, \dots, X_t = \alpha_{i_t}$ — локаторы ошибок; $Y_1 = e_{i_1}, Y_2 = e_{i_2}, \dots, Y_t = e_{i_t}$ — значения ошибок. Обозначим $Z_j = Y_j w_{i_j}$, $j = 1, \dots, t$. Тогда

$$\begin{aligned} S_i &= Z_1 X_1^i + \dots + Z_t X_t^i, \quad i = 0, 1, \dots, 2t-1; \\ S(x) &= S_0 + S_1 x + \dots + S_{2t-1} x^{2t-1}. \end{aligned}$$

Определим многочлен локаторов ошибок:

$$\sigma(x) = (1 - X_1 x)(1 - X_2 x) \cdots (1 - X_t x) = \sigma_0 + \sigma_1 x + \sigma_2 x^2 + \dots + \sigma_t x^t,$$

где $\sigma_0 = 1$. Выведем ключевое уравнение для обобщенных кодов РС. Для этого определим два многочлена $\omega(x)$ и $\Phi(x)$:

$$\omega(x) = \sum_{j=1}^t Z_j \prod_{s \neq j} (1 - X_s x); \quad \Phi(x) = \sum_{j=1}^t Z_j X_j^{2t} \prod_{s \neq j} (1 - X_s x),$$

где $\omega(x)$ называется многочленом значений ошибок для обобщенных кодов РС. Тогда

$$\begin{aligned} S(x) &= \sum_{i=0}^{2t-1} S_i x^i = \sum_{i=0}^{2t-1} \sum_{j=1}^t Z_j X_j^i x^i = \sum_{j=1}^t Z_j \sum_{i=0}^{2t-1} (X_j x)^i = \\ &= \sum_{j=1}^t Z_j \frac{1 - (X_j x)^{2t}}{1 - X_j x} = \sum_{j=1}^t \frac{Z_j}{1 - X_j x} - x^{2t} \sum_{j=1}^t \frac{X_j^{2t}}{1 - X_j x} = \frac{\omega(x)}{\sigma(x)} - x^{2t} \frac{\Phi(x)}{\sigma(x)}. \end{aligned}$$

Таким образом, получаем ключевое уравнение:

$$S(x)\sigma(x) \equiv \omega(x) \pmod{x^{2t}}. \quad (2)$$

Замечание 1. Вернемся к вопросу о наличии в векторе α нулевой компоненты. В предыдущем алгоритме этот факт не имел значения. Здесь же нужно это учитывать. Предположим, что $\alpha_i = 0$, и пусть на i -й позиции кодового вектора произошла ошибка. Так как $X_0 = \alpha_i = 0$, то данный локатор ошибки не оказывает влияния на многочлены $\sigma(x)$ и $\omega(x)$. С помощью данных многочленов можно найти только ненулевые локаторы ошибок и соответствующие им значения ошибок. Поэтому в самом конце следующего алгоритма после нахождения всех ошибок, соответствующих ненулевым локаторам, необходимо проверить, была ли еще одна ошибка на i -й позиции. Пусть \tilde{u} — вектор, в котором исправлены все ошибки, соответствующие ненулевым локаторам. Так как i -й столбец матрицы H имеет только одно ненулевое значение на первой позиции, равное w_i , то необходимо найти значение Z_0 , равное скалярному произведению вектора \tilde{u} на первую строку матрицы H . Если $Z_0 = 0$, то на i -й позиции ошибок не было. В противном случае значение ошибки на i -й позиции равно $Y_0 = Z_0 w_i^{-1}$.

Алгоритм 2 (декодирование обобщенного кода Рида–Соломона на основе алгоритма Сугиямы)

Вход: принятый вектор v .

Выход: исходный кодовый вектор u , если произошло не более $[(d - 1)/2]$ ошибок.

1. Определяется $t = [(d - 1)/2]$. Находятся первые $2t$ компонент $S_0, S_1, \dots, S_{2t-1}$ синдромного вектора vH^T . Если они все равны нулю, то полагается, что ошибок нет и процедура окончена. Полученному вектору ставится в соответствие синдромный многочлен $S(x) = \sum_{i=0}^{2t-1} S_i x^i$.

2. Пусть $r_{-1}(x) = x^{2t}$, $r_0(x) = S(x)$, $v_{-1}(x) = 0$, $v_0(x) = 1$. С помощью обобщенного алгоритма Евклида производится последовательность вычислений ($i \geq 1$):

$$r_{i-2}(x) = r_{i-1}(x)q_{i-1}(x) + r_i(x); \quad v_i(x) = v_{i-2}(x) - v_{i-1}(x)q_{i-1}(x).$$

Процесс прекращается, как только для некоторого $r_j(x)$ будет выполнено

$$\deg r_{j-1}(x) \geq t; \quad \deg r_j(x) \leq t - 1. \quad (3)$$

Тогда $\sigma(x) = \lambda v_j(x)$, $\omega(x) = \lambda r_j(x)$, где константа $\lambda \in \text{GF}(q)$ задается так, чтобы удовлетворялось условие $\sigma(0) = 1$.

3. Пусть $s = \deg \sigma(x)$. Отыскиваются s корней многочлена $\sigma(x)$ последовательной подстановкой в него ненулевых элементов поля $\text{GF}(q)$. При этом локаторы ошибок — это величины, обратные корням многочлена $\sigma(x)$.

4. Находятся Z_1, \dots, Z_s , например с помощью алгоритма Форни:

$$Z_i = \frac{\omega(X_i^{-1})}{\prod_{j \neq i} (1 - X_j X_i^{-1})}, \quad i = 1, \dots, s. \quad (4)$$

После этого находятся значения ошибок $Y_j = Z_j/w_{i_j}$, $j = 1, \dots, s$. У вектора v из i_j -го символа, $X_j = \alpha_{i_j}$, вычитается значение \tilde{Y}_j , $j = 1, \dots, s$. При этом получается вектор \tilde{u} .

Если $\alpha_i = 0$ для некоторого i , то вычисляется значение Z_0 , равное скалярному произведению вектора \tilde{u} на первую строку матрицы H . Если $Z_0 \neq 0$, то вычисляется значение ошибки $Y_0 = Z_0/w_i$. Осталось в векторе \tilde{u} из i -го символа вычесть Y_0 .

Теорема 3. Пусть $d \geq 2t + 1$, $v_j(x)$ и $r_j(x)$ — многочлены из обобщенного алгоритма Евклида с условием (3). Тогда найдется такая ненулевая константа $\lambda \in F$, для которой $\sigma(x) = \lambda v_j(x)$ и $\omega(x) = \lambda r_j(x)$.

Доказательство аналогичной теоремы для кодов РС можно найти, например, в [3].

Пример 2. Продолжим разбор примера 1, в котором рассматривался код $\text{GRS}_3(\alpha, y)$. Порождающая и проверочная матрицы кода $\text{GRS}_3(\alpha, y)$ будут иметь вид:

$$G = V_3(\alpha)\text{Diag}(y) = \begin{pmatrix} 2 & 1 & 3 & 1 & 4 & 1 & 5 \\ 0 & 1 & 6 & 3 & 2 & 5 & 2 \\ 0 & 1 & 5 & 2 & 1 & 4 & 5 \end{pmatrix};$$

$$H = V_4(\alpha)\text{Diag}(-w) = \begin{pmatrix} 4 & 1 & 5 & 1 & 2 & 1 & 3 \\ 0 & 1 & 3 & 3 & 1 & 5 & 4 \\ 0 & 1 & 6 & 2 & 4 & 4 & 3 \\ 0 & 1 & 5 & 6 & 2 & 6 & 4 \end{pmatrix},$$

где, учитывая теорему 2, $w_i = -y_i^{-1}$; $V_3(\alpha)$ и $V_4(\alpha)$ — матрицы, полученные соответственно из первых трех и четырех строк матрицы Вандермонда $V(\alpha)$. Здесь учтен тот факт, что $\prod_{j \neq i} (\alpha_j - \alpha_i) = -1$ для любого $\alpha_i \in \text{GF}(7)$.

Пусть на приемном конце получен тот же вектор $v = (2, 0, 3, 2, 3, 5, 2)$. Применим алгоритм декодирования 2.

1. Полагаем $t = [(d - 1)/2] = 2$. Вычисляем синдромный вектор $(S_0, S_1, S_2, S_3) = vH^T = (0, 2, 4, 1)$. Тогда $S(x) = 2x + 4x^2 + x^3$.

2. Полагаем $r_{-1}(x) = x^4$, $r_0(x) = S(x)$, $v_{-1}(x) = 0$, $v_0(x) = 1$. Применяем неполный обобщенный алгоритм Евклида:

$$r_{-1}(x) = r_0(x)(3 + x) + x, \quad r_1(x) = x;$$

$$v_1(x) = -(3 + x) = 4 + 6x.$$

После первого шага останавливаемся, так как $t = 2$, $\deg r_0(x) = 3$, $\deg r_1(x) = 1 < t$. Поэтому $\sigma(x) = \lambda v_1(x)$, $\omega(x) = \lambda r_1(x)$. При $\lambda = 2$ получаем $\sigma(0) = 1$, поэтому $\sigma(x) = 1 + 5x$, $\omega(x) = 2x$.

3. Корнем $\sigma(x)$ является $x_1 = 4$, поэтому $X_1 = x_1^{-1} = 2 = \alpha_2$, т. е. ошибка произошла на 2-й позиции.

4. Так как $S_1 = Z_1 X_1$, то $Z_1 = S_1 X_1^{-1} = 1$, поэтому значение ошибки равно $Y_1 = Z_1 w_2^{-1} = 1 \cdot 5^{-1} = 3$. Получаем:

$$\tilde{u} = (2, 0, 3, 2, 3, 5, 2) - (0, 0, 3, 0, 0, 0, 0) = (2, 0, 0, 2, 3, 5, 2).$$

Осталось проверить, была ли ошибка на 0-й позиции, которая соответствует элементу $\alpha_0 = 0$. Вычисляя скалярное произведение вектора \tilde{u} на первую строку матрицы H , получаем $Z_0 = 6$. Это означает, что на 0-й позиции имеется ошибка со значением $Y_0 = Z_0 w_0^{-1} = 5$. Окончательно

$$u = \tilde{u} - (5, 0, 0, 0, 0, 0, 0) = (4, 0, 0, 2, 3, 5, 2).$$

4 Декодирование обобщенных кодов Рида–Соломона на основе алгоритма Берлекэмпа–Месси (алгоритма Питерсона–Горенстейна–Цирлера)

Рассмотрим сравнение (2). Так как $\deg S(x)\sigma(x) \leq 3t - 1$, $\deg \omega(x) \leq t - 1$, то сравнение (2) имеет место, если коэффициент многочлена $S(x)\sigma(x)$ при x^i , $i = t, t+1, \dots, 2t-1$, равен нулю. Учитывая, что $\sigma_0 = 0$, получаем систему уравнений:

$$\begin{aligned} \sigma_1 S_{t-1} + \sigma_2 S_{t-2} + \cdots + \sigma_t S_0 &= -S_t; \\ \sigma_1 S_t + \sigma_2 S_{t-1} + \cdots + \sigma_t S_1 &= -S_{t+1}; \\ &\vdots \\ \sigma_1 S_{2t-2} + \sigma_2 S_{2t-3} + \cdots + \sigma_t S_{t-1} &= -S_{2t-1}. \end{aligned}$$

Запишем данную систему в матричном виде:

$$\begin{pmatrix} S_0 & S_1 & \cdots & S_{t-2} & S_{t-1} \\ S_1 & S_2 & \cdots & S_{t-1} & S_t \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{t-1} & S_t & \cdots & S_{2t-3} & S_{2t-2} \end{pmatrix} \begin{pmatrix} \sigma_t \\ \sigma_{t-1} \\ \vdots \\ \sigma_1 \end{pmatrix} = \begin{pmatrix} -S_t \\ -S_{t+1} \\ \vdots \\ -S_{2t-1} \end{pmatrix}. \quad (5)$$

Матрицу данной системы обозначим через M_t .

Теорема 4. *Матрица M_t невырождена тогда и только тогда, когда произошло t ошибок.*

Доказательство. Пусть $A = V(X_1, \dots, X_t)$ — матрица Вандермонда, $B = \text{Diag}(Z_1, \dots, Z_t)$ — диагональная матрица. Тогда доказательство следует из равенства $M_t = ABA^T$ и критерия равенства нулю определителя Вандермонда. \square

Следствие 1. Система уравнений (5) имеет единственное решение тогда и только тогда, когда произошло t ошибок.

Замечание 2. Учитывая теорему 4, осталось учесть вопрос о наличии в векторе α нулевой компоненты, если таковая имеется. Пусть $\alpha_i = 0$ (т. е. такая компонента есть). Пусть $|M_t| \neq 0$, т. е. произошло t ошибок. Если $\deg \sigma(x) = t$, то на i -й позиции ошибок не было. Если же $\deg \sigma(x) < t$, то на i -й позиции произошла ошибка. В этом случае пусть Z_0 — значение скалярного произведения вектора \tilde{u} на первую строку матрицы H , где \tilde{u} — вектор, в котором уже исправлены $t - 1$ ошибок. Тогда значение ошибки на i -й позиции равно $Y_0 = Z_0/w_i$.

Алгоритм 3 (декодирование обобщенного кода Рида–Соломона на основе алгоритма Берлекэмпа–Месси)

Вход: принятый вектор v .

Выход: исходный кодовый вектор u , если произошло не более $[(d - 1)/2]$ ошибок.

1. Определяется $t = [(d - 1)/2]$. Находятся первые $2t$ компонент $S_0, S_1, \dots, S_{2t-1}$ синдромного вектора vH^T . Если они все равны нулю, то полагается, что ошибок нет и процедура окончена.

2. Цикл: пока $|M_t| = 0$, переопределяется $t := t - 1$.

Находятся $\sigma_1, \dots, \sigma_t$ — решение системы (5). Это можно сделать с помощью алгоритма Берлекэмпа–Месси (или методом Гаусса). После этого составляется многочлен $\sigma(x)$. Пусть $s = \deg \sigma(x)$.

3. Отыскиваются s корней многочлена $\sigma(x)$ последовательной подстановкой в него ненулевых элементов поля $\text{GF}(q)$. При этом локаторы ошибок — это величины, обратные корням многочлена $\sigma(x)$.

4. Находятся Z_1, \dots, Z_s , например с помощью алгоритма Форни (4). После этого находятся значения ошибок $Y_j = Z_j/w_{i_j}$, $j = 1, \dots, s$. У вектора v из i_j -го символа, $X_j = \alpha_{i_j}$, вычитается значение Y_j , $j = 1, \dots, s$. При этом получается вектор \tilde{u} .

Если $\alpha_i = 0$ для некоторого i и $\deg \sigma(x) < t$, то вычисляется значение Z_0 , равное скалярному произведению вектора \tilde{u} на первую строку матрицы H . Вычисляется значение ошибки $Y_0 = Z_0/w_i$. Осталось в векторе \tilde{u} из i -го символа вычесть Y_0 .

Пример 3. Продолжим рассмотрение предыдущего примера. Определяем $t = [(d - 1)/2] = 2$. Составляем матрицу системы (5):

$$\left(\begin{array}{cc|c} S_0 & S_1 & -S_2 \\ S_1 & S_2 & -S_3 \end{array} \right) = \left(\begin{array}{cc|c} 0 & 2 & 3 \\ 2 & 4 & 6 \end{array} \right).$$

Так как $|M_2| \neq 0$, то полученный вектор v содержит две ошибки. Находим значения $\sigma_1 = 5, \sigma_2 = 0$. Поэтому $\sigma(x) = 1 + 5x$, из чего следует, что $X_1 = 2 = \alpha_2$. Дальше находим вектор \tilde{u} и значение ошибки на 0-й позиции $Y_0 = Z_0/w_0$, так как $\deg \sigma(x) < 2 = t$.

5 Заключение

Для декодирования кодов Гоппы (на основе которых построены некоторые перспективные постквантовые криптосистемы) хорошо известен алгоритм Паттерсона. Но он применим только для двоичных неприводимых кодов Гоппы. В то же время любой алгоритм декодирования обобщенных кодов РС можно применить и для кодов Гоппы над любым полем. Поэтому рассмотренные в данной работе алгоритмы декодирования можно применять и для кодов Гоппы [9].

Литература

1. Shiozaki A. Decoding of redundant residue polynomial codes using Euclid's algorithm // IEEE T. Inform. Theory, 1988. Vol. IT-34. Iss. 5. P. 1351–1354.
2. Gao S. A new algorithm for decoding Reed–Solomon codes // Communications, information and network security / Eds. V. Bhargava, H. V. Poor, V. Tarokh, S. Yoon. — Norwell, MA, USA: Kluwer, 2002. Vol. 712. P. 55–68.
3. Блейхут Р. Теория и практика кодов, контролирующих ошибки / Пер. с англ. — М.: Мир, 1986. 576 с. (Blahut R. E. Theory and practice of error control codes. — Addison-Wesley Pub. Co., 1983. 500 p.)
4. Elia M., Viterbo E., Bertinetti G. Decoding of binary separable Goppa codes using Berlekamp–Massey algorithm // Electron. Lett., 1999. Vol. 35. Iss. 20. P. 1720–1721.
5. Huffman W. C., Pless V. Fundamentals of error-correcting codes. — Cambridge University Press, 2003. 646 p.
6. Wang Yongge. Decoding generalized Reed–Solomon codes and its application to RLCE encryption scheme. ArXiv.org, 2017. arXiv:1702.07737 [cs.IT].
7. Status report on the first round of the NIST post-quantum cryptography standardization process: Internal Report 8240. — National Institute of Standards and Technology, 2019. 27 p. doi: 10.6028/NIST.IR.8240.
8. Федоренко С. В. Простой алгоритм декодирования алгебраических кодов // Информационно-управляющие системы, 2008. № 3. С. 23–27.
9. Ращев С. М. Об алгоритмах декодирования кодов Гоппы // Челяб. физ.-матем. ж., 2020. Т. 5. № 3. С. 327–341.

Поступила в редакцию 22.04.20

ON DECODING ALGORITHMS FOR GENERALIZED REED–SOLOMON CODES

S. M. Ratseev¹ and O. I. Cherevatenko²

¹Ulyanovsk State University, 42 Lev Tolstoy Str., Ulyanovsk 432017, Russian Federation

²Ilya Ulyanov State Pedagogical University, 4/5 Lenina Sq., Ulyanovsk 432071, Russian Federation

Abstract: The paper is devoted to decoding algorithms for generalized Reed–Solomon codes that are based on algorithms for Reed–Solomon codes. The Gao, Sugiyama, and Berlekamp–Massey algorithms (Peterson–Gorenstein–Zierler algorithm) are given. The first of these algorithms belongs to syndrome-free decoding algorithms, the others — to syndrome decoding algorithms. The relevance of these algorithms is that they are applicable for decoding Goppa codes which are the basis of some promising postquantum cryptosystems. These algorithms are applicable for Goppa codes over an arbitrary field as opposed to the well-known Patterson decoding algorithm for binary Goppa codes.

Keywords: error-correcting codes; Reed–Solomon codes; Goppa codes; code decoding

DOI: 10.14357/08696527200408

References

1. Shiozaki, A. 1988. Decoding of redundant residue polynomial codes using Euclid's algorithm. *IEEE T. Inform. Theory* IT-34(5):1351–1354.
2. Gao, S. 2002. A new algorithm for decoding Reed–Solomon codes. *Communications, information and network security*. Eds. V. Bhargava, H. V. Poor, V. Tarokh, and S. Yoon. Norwell, MA: Kluwer. 712:55–68.
3. Blahut, R. E. 1983. *Theory and practice of error control codes*. Addison-Wesley Pub. Co. 500 p.
4. Elia, M., E. Viterbo, and G. Bertinetti. 1999. Decoding of binary separable Goppa codes using Berlekamp–Massey algorithm. *Electron. Lett.* 35(20):1720–1721.
5. Huffman, W. C., and V. Pless. 2003. *Fundamentals of error-correcting codes*. Cambridge University Press. 646 p.
6. Wang, Yongge. 2017. Decoding generalized Reed–Solomon codes and its application to RLCE encryption scheme. arXiv.org. Available at: <https://arxiv.org/abs/1702.07737> (accessed October 13, 2020).
7. NISTIR 8240. 2019. Status report on the first round of the NIST post-quantum cryptography standardization process. National Institute of Standards and Technology. 27 p. doi: 10.6028/NIST.IR.8240.
8. Fedorenko, S. V. 2008. Prostoy algoritm dekodirovaniya algebraicheskikh kodov [A simple algorithm for decoding algebraic codes]. *Informatsionno-upravlyayushchie sistemy* [Information and Control Systems] 3:23–27.

9. Ratseev, S. M. 2020. Ob algoritmakh dekodirovaniya kodov Goppy [On decoding algorithms for Goppa codes]. *Chelyab. fiz.-matem. zh.* [Chelyabinsk Physical and Mathematical J.] 5(3):327–341.

Received April 22, 2020

Contributors

Ratseev Sergey M. (b. 1979) — Doctor of Science in physics and mathematics, associate professor, Department of Information Security and Control Theory, Ulyanovsk State University, 42 Lev Tolstoy Str., Ulyanovsk 432017, Russian Federation; ratseevsm@mail.ru

Cherevatenko Olga I. (b. 1983) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Higher Mathematics, Ilya Ulyanov State Pedagogical University, 4/5 Lenina Sq., Ulyanovsk 432071, Russian Federation; chai@pisem.net

РАЗВИТИЕ ГИБРИДНОЙ МНОГОЯДЕРНОЙ РЕКУРРЕНТНОЙ АРХИТЕКТУРЫ НА ПЛИС*

*Ю. А. Степченков¹, Н. В. Морозов², Ю. Г. Дьяченко³, Д. В. Хилько⁴,
Д. Ю. Степченков⁵*

Аннотация: Представлен результат модификации гибридной многоядерной архитектуры рекуррентного сигнального процессора (ГМАРСП) и ее апробации в виде макетного образца на отладочной плате нового поколения HAN Pilot Platform с программируемой логической интегральной схемой (ПЛИС) Intel Arria10 SoC 10AS066K3F40E2SG на основе VHDL (very high speed integrated circuits) модели уровня регистровых передач. Гибридная многоядерная архитектура рекуррентного сигнального процессора состоит из ведущего фон-неймановского процессора на управляющем уровне и потокового процессора с восемью вычислительными ядрами (ВЯ) на операционном уровне. Вычислительные ядра объединены капсулальным распределителем, обеспечивающим развертывание алгоритмической капсулы в параллельно-последовательный поток команд, и работают с 32-разрядными данными. Аппаратная реализация двухъядерного процессора управляющего уровня Cortex-A9 обеспечила существенное повышение производительности всей ГМАРСП и увеличение точности обработки данных за счет использования 32-разрядных данных с фиксированной точкой. Апробация VHDL-модели модифицированной ГМАРСП на типовом приложении цифровой обработки данных — распознавателе изолированных слов (РИС) — показала ее высокую эффективность при работе в режиме реального времени.

Ключевые слова: рекуррентный сигнальный процессор; гибридная многоядерная архитектура; VHDL-модель; ПЛИС; распознаватель изолированных слов

DOI: 10.14357/08696527200409

*Исследование выполнено при поддержке Российского научного фонда (проект 19-11-00334).

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, YStepchenkov@ipiran.ru

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, diaura@mail.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dhilko@yandex.ru

⁵Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, stepchenkov@mail.ru

1 Введение

Потоковая (data-flow) архитектура стала в настоящее время наиболее активно развивающимся направлением в разработке систем параллельных вычислений. В Институте проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук в течение ряда лет разрабатывается ГМАРСП [1, 2].

Тестовая отладка ГМАРСП проводились на основе ее VHDL-модели уровня регистровых передач и программируемой логической интегральной схемы (ПЛИС) 5CGTFD9E5F35C7 (Intel) в составе отладочной платы «Cyclone V GT FPGA Development Kit» (Terasic). С помощью системы Quartus (Intel) VHDL-модель ГМАРСП синтезировалась и загружалась в ПЛИС. В качестве предметной области для отладки архитектуры ГМАРСП был выбран РИС [3].

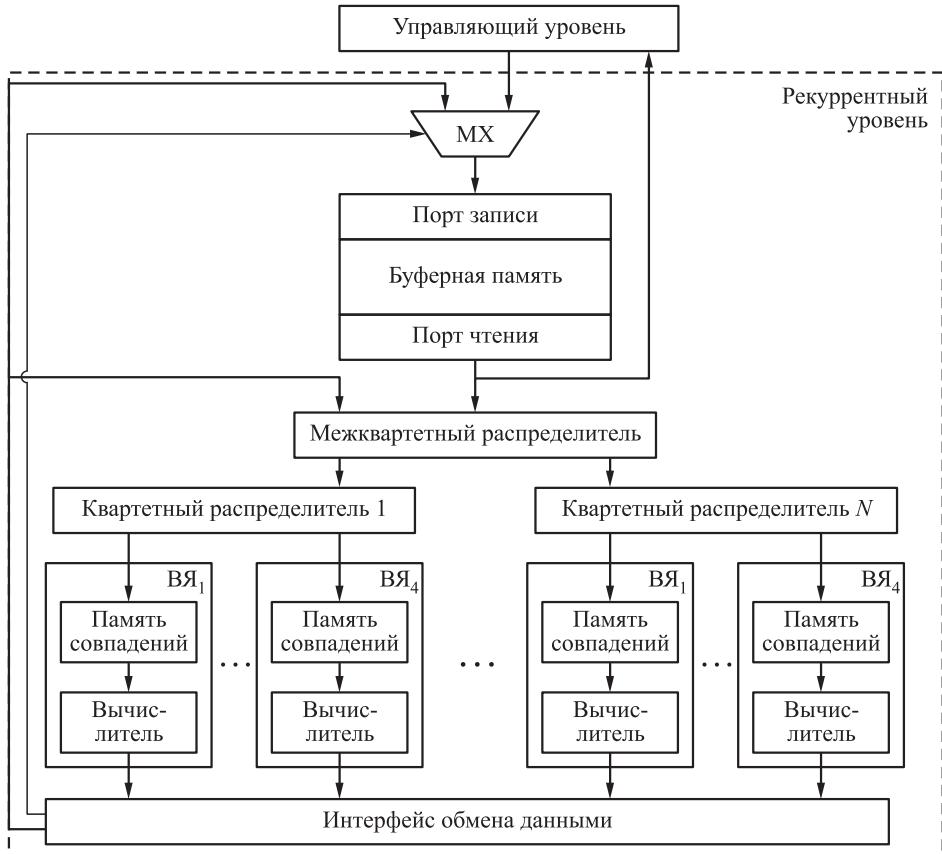
Для верификации ГМАРСП была использована существующая реализация алгоритмов РИС на имитационной модели архитектуры. Верификация проводилась по факту побитового сравнения результатов исполнения алгоритмов как на имитационной и аппаратной моделях, так и на ПЛИС-реализациях. Полученные результаты подтвердили корректность реализации.

Тестовая отладка проекта РИС подтвердила высокую эффективность ГМАРСП, реализованной в виде VHDL-модели, при выполнении различных алгоритмов цифровой обработки речевых сигналов [4]. Однако ограниченные аппаратные ресурсы платы «Cyclone V GT» заставили существенно ограничить функциональные возможности ГМАРСП для обеспечения работоспособности РИС в ПЛИС. Реальная частота работы ГМАРСП оказалась невысокой (12,5 МГц) и затруднила реализацию в режиме реального времени сложных алгоритмов цифровой обработки данных. Кроме того, имеющихся на плате ресурсов оказалось недостаточно для расширения разрядности и увеличения числа вычислительных ядер в ГМАРСП.

Данная статья представляет отладочную плату нового поколения (Arria 10 SoC Development Kit с установленной на ней ПЛИС Arria10 SoC 10AS066N3F40E2SG1 (Intel)) и оценивает целесообразность ее использования для разработки и отладки ГМАРСП, реализованной в виде VHDL-модели уровня регистровых передач.

2 Расширение структуры гибридной многоядерной архитектуры рекуррентного сигнального процессора

Структурная схема расширения ГМАРСП показана на рисунке. Двухпортовая буферная память обеспечивает максимальную производительность при обмене данными с процессором управляющего уровня, загрузке операндами рекуррентного уровня и обновлении капсул результатами работы рекуррентного уровня. Межквартетный распределитель формирует операнды для квартетных



Структурная схема ГМАРСП

распределителей, каждый из которых, в свою очередь, формирует пакеты («горсти») для своих четырех ВЯ, результаты работы которых интерфейс обмена данными отправляет обратно в межквартетный распределитель или записывает в буферную память.

В сравнении с ГМАРСП, описанной в [4], структура нового поколения ГМАРСП, представленная на рисунке, содержит больше вычислительных ядер и работает с большей разрядностью обрабатываемых данных (32 вместо 16). Это обеспечивает повышение точности вычислительных операций и более эффективное выполнение распараллелиемых операций. Гибридная многоядерная архитектура рекуррентного сигнального процессора нового поколения также реализована в виде VHDL-модели уровня регистровых передач.

3 ПЛИС для отладки расширенной структуры гибридной многоядерной архитектуры рекуррентного сигнального процессора

Выбор платы и ПЛИС на ней определяется ожидаемой сложностью ГМАРСП и проектов на ее основе. Увеличение числа вычислительных ядер и расширение разрядности внутренних шин данных требуют значительных аппаратных ресурсов и возможностей «фабрики межсоединений».

Технологический уровень изготовления ПЛИС свидетельствует о ее потенциальной производительности и характеристиках «фабрики межсоединений». Чем меньше технологические нормы, тем более высокую производительность может обеспечить ПЛИС за счет уменьшения задержек переключения логических элементов и повышения пропускной способности (трассируемости) «фабрики межсоединений».

Таблица 1 Характеристики ПЛИС в составе отладочных плат

ПЛИС	Логиче- ские эле- менты, тыс.	Адап- тивные логи- ческие модули	Блоки памя- ти M10K	Блоки памя- ти M20K	Регист- ры (триг- геры)	Гло- баль- ные такто- вые цепи	Умно- жи- тели	Аппа- ратный про- цессор	Техно- логиче- ские нор- мы, нм
5CGTFD9E5F35C7	301	113 560	1220	—	454 240	16	684	—	28
10AS066N3F40E2SG1	660	250 540	—	2133	1 002 160	32	1688	ARM Cortex-A9, 1,5 ГГц	20

Анализ возможных вариантов ПЛИС и отладочных плат показал, что наиболее предпочтительный вариант — плата «Arria 10 SoC Development Kit» с установленной на ней ПЛИС «Arria10 SoC 10AS066N3F40E2SG1» (Intel). Таблица 1 показывает ее характеристики в сравнении с ПЛИС 5CGTFD9E5F35C7. Аппаратные ресурсы ПЛИС 10AS066N3F40E2SG1 по всем позициям более чем в 2 раза превосходят соответствующие показатели ПЛИС 5CGTFD9E5F35C7. Меньшие технологические нормы гарантируют увеличение трассировочной способности «фабрики межсоединений», которая облегчает размещение и разводку компонентов проекта. Наличие аппаратной реализации процессора ARM Cortex-A9, выполняющего функции управляющего уровня в расширенной ГМАРСП, обеспечивает дополнительное повышение ее производительности.

4 Характеристики реализации расширенной гибридной многоядерной архитектуры рекуррентного сигнального процессора

Экспериментальные результаты реализации исходного варианта ГМАРСП, реализованной в виде VHDL-модели уровня регистровых передач, на отладочной плате «Arria 10 SoC Development Kit» с ПЛИС «Arria10 SoC

Таблица 2 Максимально возможная частота синхронизации ГМАРСП в разных ПЛИС

Реализация РИС на ПЛИС	Управляющий уровень, МГц		Рекуррентный уровень, МГц
	Программная реализация	Аппаратная реализация	
5CGTFD9E5F35C7 исходная	90	—	12,5
10AS066N3F40E2SG1 исходная расширенная	160	1500	26,0
	160	1500	25,2

10AS066N3F40E2SG1» показали, что ее аппаратные ресурсы обеспечили автоматическое повышение производительности ГМАРСП в режиме реального времени даже без использования встроенного процессора ARM Cortex-A9. В таблице 2 приведены характеристики исходного проекта ГМАРСП в двух ПЛИС. Новая ПЛИС обеспечивает как минимум удвоение производительности исходного варианта ГМАРСП.

Реализация РИС на расширенном варианте ГМАРСП привела к незначительному ухудшению производительности (см. последнюю строку табл. 2). При этом оказались задействованы лишь 40% аппаратных ресурсов ПЛИС 10AS066N3F40E2SG1.

Таким образом, использование ПЛИС 10AS066N3F40E2SG1 обеспечивает запуск РИС на исходном варианте ГМАРСП в режиме реального времени и делает возможным расширение разрядности обрабатываемых в ГМАРСП данных до 64 бит и увеличение числа вычислительных ядер до 16.

5 Заключение

Гибридная многоядерная архитектура рекуррентного сигнального процессора нового поколения, реализованная в виде VHDL-модели уровня регистровых передач, является ресурсоемкой за счет распараллеливания вычислительных задач и использования FIFO (first in, first out) в разных блоках архитектуры.

ПЛИС 10AS066N3F40E2SG1 обеспечивает удвоение производительности исходного варианта ГМАРСП и запуск речевых приложений в режиме реального времени. Она гарантирует расширение разрядности данных в ГМАРСП до 32 бит и увеличение числа вычислительных ядер, а в перспективе делает возможной обработку 64-разрядных данных.

Литература

- Stepchenkov Yu., Khilko D., Diachenko Yu., Shikunov Yu., Shikunov D. Software and hardware testing of dataflow recurrent digital signal processor // IEEE East-West Design & Test Symposium Proceedings. — IEEE, 2016. P. 168–171.

2. *Shikunov Yu., Stepchenkov Yu., Khilko D.* Recurrent mechanism developments in the data-flow computer architecture // IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conference Proceedings. — IEEE, 2018. P. 1413–1418.
3. Степченков Ю. А., Дьяченко Ю. Г., Хилько Д. В., Петрухин В. С. Рекуррентная потоковая архитектура: особенности и проблемы реализации // Проблемы разработки перспективных микро- и наноэлектронных систем / Под общ. ред. акад. РАН А. Л. Степпиковского. — М.: ИППМ РАН, 2016. Ч. 2. С. 120–127.
4. *Stepchenkov Yu., Morozov N., Khilko D., Shikunov Yu., Orlov G.* Hybrid multicore recurrent architecture approbation on FPGA // IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conference Proceedings.— IEEE, 2019. P. 1075–1078.

Поступила в редакцию 25.03.20

MULTICORE HYBRID RECURRENT ARCHITECTURE EXPANSION ON FPGA

*Yu. A. Stepchenkov, N. V. Morozov, Yu. G. Diachenko, D. V. Khilko,
and D. Yu. Stepchenkov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper presents the result of modification of the multicore hybrid architecture for recurrent signal processing (HARSP) and discusses its approbation as a prototype on the next-generation HAN Pilot Platform development board with FPGA (field-programmable gate array) Intel Arria10 SoC 10AS066K3F40E2SG on the basis of the register transfer level VHDL (very high speed integrated circuits) model. Hybrid architecture for recurrent signal processing contains the control level, implemented as von Neumann processor, and the operational level represented by the data-flow processor with eight computing cores. A capsule distributor combines all computing cores. It provides algorithmic capsule explication into a parallel-serial command flow and processes 32-bit data. Hardware implementation of the control level dual-core processor Cortex-A9 improved HARSP performance radically and increased data processing accuracy due to using 32-bit fixed-point operands. Modified HARSP VHDL-model approbation on a typical data processing application, namely, isolated word recognition, proved HARSP high efficiency in real-time mode operation.

Keywords: recurrent signal processor; multicore hybrid architecture; data-flow; VHDL-model; FPGA; development board; isolated word recognizer

DOI: 10.14357/08696527200409

Acknowledgments

The research was funded by a grant from the Russian Science Foundation (project No. 19-11-00334).

References

1. Stepchenkov, Yu., D. Khilko, Yu. Diachenko, Yu. Shikunov, and D. Shikunov. 2016. Software and hardware testing of dataflow recurrent digital signal processor. *IEEE East-West Design & Test Symposium Proceedings*. IEEE. 168–171.
2. Shikunov, Yu., Yu. Stepchenkov, and D. Khilko. 2018. Recurrent mechanism developments in the data-flow computer architecture. *IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conference Proceedings*. IEEE. 1413–1418.
3. Stepchenkov Yu. A., Yu. G. Diachenko, D. V. Khilko, and V. S. Petrukhin. 2017. Rekurrentnaya potokovaya arkhitektura: osobennosti i problemy realizatsii [Recurrent data-flow architecture: Features and realization problems]. *Problemy razrabotki perspektivnykh mikro- i nanoelektronnykh system* [Problems of perspective micro- and nanoelectronic systems development]. Moscow: IDPM RAS. 2:120–127.
4. Stepchenkov, Yu., N. Morozov, D. Khilko, Yu. Shikunov, and G. Orlov. 2019. Hybrid multi-core recurrent architecture approbation on FPGA. *IEEE Russia Section Young Researchers in Electrical and Electronic Engineering Conference Proceedings*. IEEE. 1075–1078.

Received March 25, 2020

Contributors

Stepchenkov Yuri A. (b. 1951) — Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; YStepchenkov@ipiran.ru

Morozov Nikolai V. (b. 1956) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; NMorozov@ipiran.ru

Diachenko Yuri G. (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaura@mail.ru

Khilko Dmitri V. (b. 1987) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; dhilko@yandex.ru

Stepchenkov Dmitri Yu. (b. 1973) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; stepchenkov@mail.ru

МЕТОД СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ КВАДРОДЕРЕВА

Ю. А. Маняков¹, А. И. Сорокин²

Аннотация: Представлен метод сегментации цветного изображения на основе квадродерева, состоящий из нескольких шагов. Первый из них — обнаружение границ, основанное на трехканальном цвете и двух масках. Далее применяется алгоритм уточнения контура. Алгоритм сегментации делится на две части. В первой части изображение разделяется на максимальное число сегментов. Во второй части сегменты объединяются на основе FSM-таблицы. В статье представлены результаты сегментации цветных изображений, полученные на основе описанного алгоритма.

Ключевые слова: цветное изображение; сегментация; квадродерево; ребро; граница; истончение; уменьшение цвета; разделение; слияние; пиксель

DOI: 10.14357/08696527200410

1 Введение

Сегментация — одна из фундаментальных операций, используемых в обработке изображений и, в частности, в области компьютерного зрения [1] для извлечения необходимой информации из сцены. Несмотря на то что этому вопросу посвящено немало исследований, задача остается актуальной [2]. В общем, можно сказать, что сегментация заключается в разбиении изображения на различные непересекающиеся области. Такие области, как правило, однородны по некоторому признаку, вычисленному на основе изображения.

В основе алгоритмов сегментации лежат различные подходы [3], из которых можно выделить четыре основных.

1. Обнаружение контуров. Этот подход широко используется для сегментации черно-белых изображений [4]. В таких изображениях поиск краев осуществляется методом градиентной аппроксимации на основе вычисления первой производной (операторы Канни, Собеля и др. [5]). С точки зрения визуальной оценки они выглядят хорошо, но не соответствуют требованиям непрерывности и минимальной толщины контурных линий.

Градиентный поиск цветных изображений осуществляется, в основном, двумя способами [6]: для всех вариаций цветовых каналов используется одна

¹Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, maniakov_yuri@mail.ru

²Орловский филиал Федерального исследовательского центра «Информатика и управление» Российской академии наук, webdi@mail.ru

мера; градиент канала вычисляется отдельно и объединяет эти вычисления по определенным критериям.

2. Выращивание регионов. Этот процесс начинается с отбора исходного «семени». Точки вокруг семени группируют по некоторому критерию однородности. Для сегментации цветных изображений используется большое число методов, основанных на обработке полутоновых изображений [7]. Необходимо учитывать, что при использовании данного подхода имеет место зависимость полученных результатов от порядка обработки точек изображения. Также в этом случае необходимо проанализировать каждый пиксель и соотнести его с определенной однородной областью.
3. Методы разделения и слияния. Эта группа использует изображение в качестве начальной области для разделения [3]. После создания однородных областей деление заканчивается. Кроме того, соседние регионы связаны на основе квадратичного дерева. Связанные области должны быть гладкими и непрерывными. Для этого используют марковское случайное поле, стабилизированные уравнения обратной диффузии и др.
4. Кластеризация. В процессе кластеризации пиксели изображения делятся на группы, которые не пересекаются. Характеристики объектов, входящих в группу, должны быть близкими. В противном случае объекты, принадлежащие к разным группам, должны существенно отличаться. Соответственно, расположение пикселей в каждом кластере будет определяться цветовыми вариациями изображения. В этом случае задача кластеризации определяется как поиск определенного числа паттернов, однозначно определяющих число областей [7].

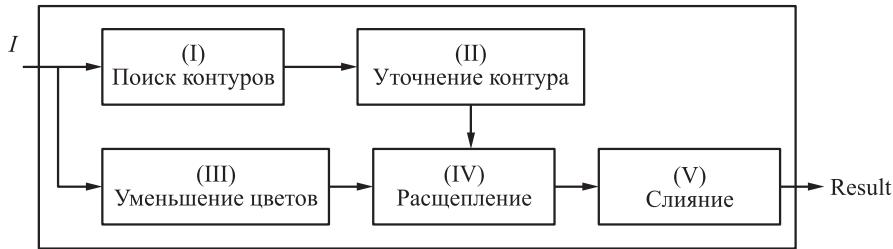
В данной работе предложен метод сегментации изображений на основе разделения и слияния. Этот метод состоит из поиска ребер; истончения вычисленных границ; уменьшения цвета; разбиения изображения квадрата на основе цвета и учета информации о границах; объединения однородных областей на основе графа. Данный алгоритм применяется для извлечения информации о сегментах и анализа их для поиска опорных точек в задачах стереозрения.

2 Сегментация

В формальном виде сегментация изображения может быть представлена следующим образом [3]. Пусть изображение I размером $W \times H$ является некоторой дискретной двумерной функцией $f(x, y)$, где I — это двумерная матрица, в которой определена точка на изображении, а соответствующие точечные элементы матрицы задают ее цветовую характеристику.

Пусть P — предикат однородности. Тогда сегментация разбивает I на множество N областей S_n , $n = 1, \dots, N$, таких что

$$(1) \bigcup_{n=1}^N S_n = I, S_n \cap S_m = 0, n \neq m;$$

**Рис. 1** Этапы сегментации

- (2) $P(S_n) = 1 \forall n;$
- (3) $P(S_n \cup S_m) = 0, \forall S_n, S_m$ — смежных.

Первое и второе условия говорят о том, что каждый пиксель изображения принадлежит только одной области [8]. В третьем указывается, что области однородны.

Процесс сегментации изображения состоит из нескольких этапов (рис. 1).

Первоначально осуществляется выделение контуров на основе цветовых значений. После этого выполняется уточнение контура. Далее уменьшается число цветов на изображении и на основе информации о цвете и о контурах выполняется расщепление, а затем слияние.

Структура данных. Как упоминалось ранее, сегментация основана на квадродержеве [9]. Изображение хранится в дереве таким образом, что каждый узел имеет четыре потомка — NW, NE, SW и SE.

Таким образом, дескриптор прямоугольной области изображения и узла дерева (Node) описывается следующим образом:

$$\text{Node} = \left\{ \text{Level}, \left\{ \begin{array}{c} \text{RegionID} \\ \text{PartitionID} \\ \text{ChNum} \end{array} \right\}, \text{avgColor}, \text{nPoint}, \text{nBorder}, \text{Rect}, \text{Env} \right\}.$$

Здесь Level — высота текущего узла квадродержева; RegionID — уникальный идентификатор текущего узла, который вычисляется следующим образом [10]:

$$\text{RegionID} = \text{PartitionId} \cdot 10 + \text{ChNum}; \quad (1)$$

PartitionID — идентификатор родителя; ChNum — номер узла в текущем разделе: NW — 1, NE — 2, SW — 3, SE — 4; avgColor — среднее значение цвета; nPoint — число точек в текущем узле; nBorder — число точек, принадлежащих контуру в текущем узле; Rect — границы узла: Left, Right, Top, Bottom; Env — идентификаторы соседних узлов: NorthID, EastID, SouthID, WestID.

Обнаружение контуров. Для обнаружения контуров используется подход, описанный в работе [11] — поиск изменения цвета с помощью масок (рис. 2). Для

$i-1, j-1$	$i-1, j$	$i-1, j+1$
$I, j-1$	i, j	$i, j+1$
$i+1, j-1$	$i+1, j$	$i+1, j+1$

(a)

0	0	0
4	0	-4
0	0	0

(б)

Рис. 2 Применение масок: (а) расположение пикселей; (б) маски 0° и 90°

снижения вычислительной нагрузки используются две маски: горизонтальная (0°) и вертикальная (90°) (рис. 2, б).

Пороговое значение вычисляется следующим образом:

$$T = (\text{dp}(i, j) | \text{Avg}) ;$$

$$\begin{aligned} & (\text{dp}(i, j) | 0^\circ) = \\ &= \sum_{i=0}^{\max I} \sum_{j=0}^{\max J} [|(K_{\max} R(i, j - 1) + G(i, j - 1) + B(i, j - 1)) a(i, j - 1) + \\ &\quad + (K_{\max} R(i, j + 1) + G(i, j + 1) + B(i, j + 1)) a(i, j + 1)|] ; \\ & (\text{dp}(i, j) | 90^\circ) = \\ &= \sum_{i=0}^{\max I} \sum_{j=0}^{\max J} [|(K_{\max} R(i - 1, j) + G(i - 1, j) + B(i - 1, j)) b(i - 1, j) + \\ &\quad + (K_{\max} R(i + 1, j) + G(i + 1, j) + B(i + 1, j)) b(i + 1, j)|] . \end{aligned}$$

Утончение контура. Для прореживания границ применяется алгоритм, описанный в [12], в котором применяются два шага итерации — один для горизонтальных и вертикальных линий, второй для наклонных. После этого необходимо провести отбор. Наиболее подходящие границы выбираются на основе компонентов связности. Учитываются непрерывность и длина границ. Результатом этой операции становится создание бинарной матрицы brdMtx размерности $W \times H$.

Прореживание цвета. Для уменьшения цвета изображения используется подход, описанный в [13]. Такой выбор обусловлен простотой реализации и качеством получаемых результатов.

Расщепление. Процесс расщепления осуществляется путем деления родительского узла на четыре дочерних. Для каждого квадранта проверяется, насколько однородна часть изображения, лежащая в его границах. В общем случае, если предикат, примененный к четырем дочерним узлам, истинен, то расщепления не происходит:

$$\text{Node}_n = \begin{cases} \bigcup_{j=1}^4 \text{Node}_{nj}, & \text{если } P(\text{Node}_n) = \text{false}; \\ \text{Node}_n & \text{иначе.} \end{cases}$$

Для каждого дочернего узла вычисляется среднее значение его цвета avgColor и число граничных пикселей numBrd. Если манхэттенское расстояние между avgColor и цветом узла pColor меньше T_{clr} или $\text{numBrd} \geq T_{\text{brd}}$, то текущий узел делится на четыре.

Процесс расщепления можно представить следующим образом:

$$P(\text{Node}_n) = \begin{cases} \text{false}, & \text{если } \sum_{x=0}^H \sum_{y=0}^W |\text{avgColor}_{\text{Node}_n}(f(x, y)) - \\ & \quad - \text{pColor}_{x, y \in \text{Node}_n}(f(x, y))| \leq T_{\text{clr}}; \\ \text{false}, & \text{если } \text{numBrd}_{\text{Node}_n} \geq T_{\text{brd}}; \\ \text{true} & \text{иначе.} \end{cases}$$

Алгоритм разделения узлов выглядит следующим образом. На вход подается изображение с уменьшенным числом цветов Iq и brdMtx. В цикле выполняются следующие действия:

- вычисление координат потомка;
- поиск среднего цвета для Node_n каждого канала;
- вычисление дистанции между avgColor_{Node_n} и pColor_{x, y ∈ Node_n};
- если dst_c ≤ T_{clr} или brd_n ≥ T_{brd}, то выполнение следующего действия, иначе конец;
- добавление информации о текущем узле, вычисление RegionID по формуле (1).

На выходе получаем $\bigcup_{n=1}^N S_n = I$.

Результатом выполнения этого алгоритма будет максимальное разбиение изображений.

Слияние. Этап объединения начинается с определения соседей для каждого узла. Эта операция основана на модификации подхода, предложенного в работе [14], т. е. используется конечный автомат (КА), представленный в таблице.

В отличие от [14] используются четыре состояния. Это делается для снижения вычислительной нагрузки.

Два значения (число и тире) в ячейке указывают на то, что поиск соседей в этом направлении должен продолжаться для следующего прохода (R, L, D и U). Одно значение в ячейке означает, что текущий квадрант последний и поиск в этом направлении необходимо остановить.

Ниже приведена последовательность шагов алгоритма поиска соседей для узла:

Конечный автомат для соседей квадрантов по четырем направлениям

Направление	Квадрант			
	1	2	3	4
R	2	1-	4	3-
L	2-	1	4-	3
D	3	4	1-	2-
U	3-	4-	1	2

- если Left, Right, Top, Bottom равны 0 или размеру изображения (W, H), то в соответствующем направлении соседей нет;
- вычислить соседа для Node_n : в цикле для всех существующих направлений (не равных 0) вычислить все значения соседних квадрантов на основе ChNum в соответствии с таблицей;
- занести значение в Env.

Конечный автомат возвращает число, обозначающее соседний квадрант в соответствующем направлении. Уровень соседа вычисляется до размера не менее размера текущего узла [14]. Но квадрант может граничить с соседом любого размера. В этом случае необходимо уточнение размера соседей. Поскольку Env содержит идентификатор соседей, необходимо провести направленный спуск по дереву и определить существование узла с соответствующим идентификатором. Ниже представлен алгоритм поиска существующих соседей.

На вход подается дерево квадрантов QTree с вычисленным Env. Далее выполняются следующие действия:

- просмотр значений из Env и занесение их в стек strID;
- в цикле осуществить просмотр значений для верхнего значения strID: занести его в текущий путь curWay, если его длина равна 0 или узел с текущим номером не существует, то закончить дальнейшее рассмотрение текущего пути и вернуть существующий путь и указатель на последний существующий узел этого пути, иначе занести значение в пройденный путь rtnWay, повторить для следующего значения;
- создать связь из текущего узла в узел с индексом в rtnWay;
- сохранить пройденный путь rtnWay и указатель на узел.

На выходе алгоритма — QTree с вычисленным существующим Env и ссылками на соседей. Результатом работы этого алгоритма будет граф, в котором имеются переходы от узла к каждому существующему соседу, не меньше текущего узла.

Для того чтобы осуществить слияние, необходимо перейти к линейным формам квадродерева QTree — LnQTree — и провести сортировку по числу точек в квадранте в порядке возрастания. Это делается для ускорения слияния.

Алгоритм слияния на вход получает дерево в отсортированном линейном виде LnQTree и предполагает следующую последовательность шагов:

- осуществить просмотр массива LnQTree, пока в нем имеются элементы;
- извлечь крайний элемент массива LnQTree и посмотреть массив слияния Segment (содержащий индекс сегмента, которому принадлежит квадрант, число точек и цвет): если он пуст, то создать первый сегмент, иначе создать новый сегмент;
- пометить узел как посещенный;
- в цикле осуществить просмотр узлов, смежных с узлами текущего сегмента, и поместить их в vNgbrSgmNode;
- вычислить близость цвета сегмента и узла соседа и, если она меньше порога, добавить этот узел в сегмент и отметить узел как посещенный;
- осуществить просмотр vNgbrSgmNode и извлечь узлы, которые были посещены;
- аналогично пройтись по массиву LnQTree и удалить посещенные узлы.

На выходе алгоритма — массив сегментов, в котором каждый сегмент состоит из узлов, близких по цветовым характеристикам.

Таким образом, каждый узел можно классифицировать и сопоставить с определенным сегментом. Принадлежность к соответствующему сегменту определяется исходя из значения, указанного в segmentnumber, используемого в качестве индекса массива сегментов.

3 Результаты

Результат работы описанного алгоритма приведен на рис. 3.

4 Выводы

Как видно из представленных результатов, алгоритм сегментации хорошо справляется с изображениями, содержащими значительное число мелких деталей, потеря которых создает трудности при анализе. При этом фрагменты изображения, представляющие собой однородную область, объединяются в более крупные сегменты.

Параметры, используемые алгоритмом при сегментации, были установлены эмпирически. Возможно, было бы целесообразно выбрать их автоматически в зависимости от типа входного изображения. По мере дальнейшей работы возможно улучшение выбора границ и, вероятно, использование дополнительных индикаторов, например основанных на текстурах, для улучшения результатов сегментации, в частности для сложных изображений.

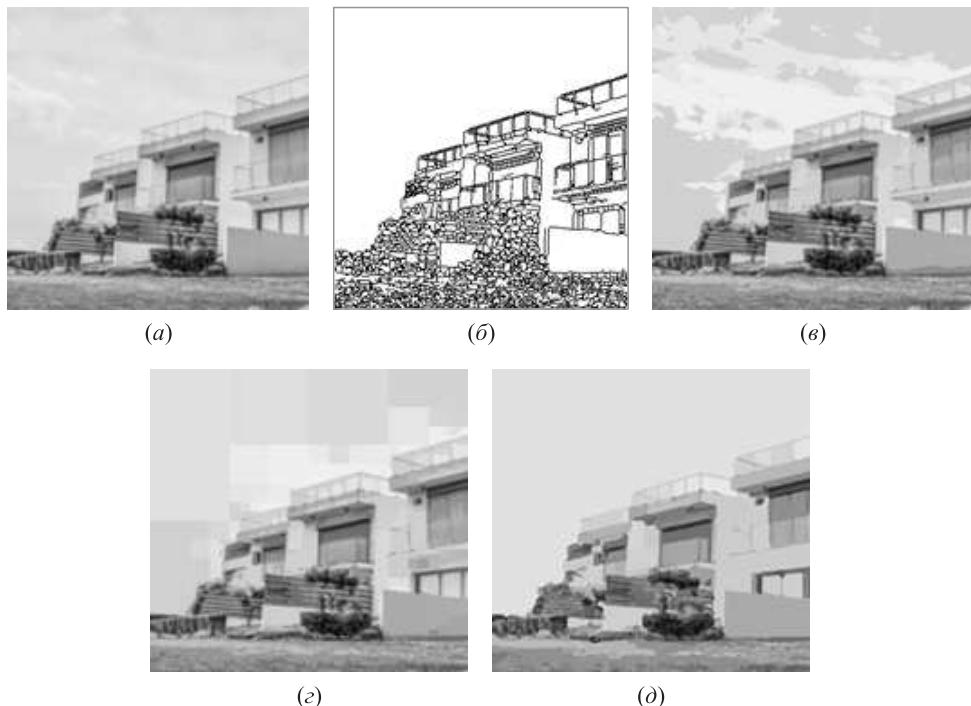


Рис. 3 Результаты работы алгоритма: (а) исходное изображение; (б) утончение контуров; (в) уменьшение цветов на изображении; (г) разделение; (д) слияние

Литература

1. Stojmenovic M., Montero A. S., Nayak A. Colour and texture based pyramidal image segmentation // Conference (International) on Audio, Language and Image Processing Proceedings. — IEEE, 2010. P. 778–786.
2. Correa-Tome F. E., Sanchez-Yanez R. E. Integral split-and-merge methodology for real-time image segmentation // J. Electron. Imaging, 2015. Vol. 24. Iss. 1. P. 1–11.
3. Волков В. Н., Сорокин А. И. Классификация методов сегментации цветных изображений // Информационные системы и технологии, 2018. № 4(108). С. 32–44.
4. Fu K. S., Mui J. K. A survey on image segmentation, pattern recognition // Pattern Recogn., 1981. Vol. 13. Iss. 1. P. 3–16.
5. Ferhat F. A., Kerdjidj O., Messaoudi K., Abdelwahhab B. Implementation of Sobel, Prewitt, Roberts edge detection on FPGA // World Congress in Computer Science, Computer Engineering, and Applied Computing Proceedings, 2013. P. 1–4.
6. Lucchese L., Mitray S. K. Color image segmentation: A state-of-the-art survey // P. Indian National Science Academy, 2001. Vol. 67. Iss. 2. P. 207–221.

7. *Verges-Llahi J.* Color constancy and image segmentation techniques for applications to mobile robotics. — Barcelona, Spain: Polytechnic University of Catalonia, 2005. PhD Thesis. 23 p.
8. *Lorenzo-Navarro J., Hernández-Tejera M.* Image segmentation using a modified split and merge technique // *Cybernet. Syst.*, 2007. Vol. 25. Iss. 1. P. 137–162.
9. *Aizawa K., Motomura K., Kimura S., Kadowaki R., Fan J.* Constant time neighbor finding in quadtrees: An experimental result // *3rd Symposium (International) on Communications, Control and Signal Processing Proceedings*. — IEEE, 2008. P. 505–510.
10. *Kelkar D., Gupta S.* Improved quadtree method for split merge image segmentation // *1st Conference (International) on Emerging Trends in Engineering and Technology Proceedings*. — IEEE, 2008. P. 44–47.
11. *Сорокин А. И., Волков В. Н. Шульдешова О. В.* Синтетический алгоритм обнаружения контуров объектов на основе RGB-координат // *Информационные системы и технологии*, 2019. № 2(112). С. 27–34.
12. *Kwon J. S., Woong J., Kang E. K.* An enhanced thinning algorithm using parallel processing // *Conference (International) on Image Processing Proceedings*. — IEEE, 2001. Vol. 3. P. 752–755.
13. *Laganière R.* OpenCV 2 computer vision application programming cookbook. — Packt Publishing, 2011. 306 p.
14. *Yoder R., Bloniarz P.* A practical algorithm for computing neighbors in quadtrees, octrees, and hyperoctrees // *Conference (International) on Modeling, Simulation & Visualization Methods Proceedings*. — CSREA Press, 2006. P. 249–255.

Поступила в редакцию 14.09.20

QUADTREE BASED COLOR IMAGE SEGMENTATION METHOD

Yu. A. Maniakov and A. I. Sorokin

Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation

Abstract: The paper presents a color image segmentation method and an algorithm based on quadtree. The proposed method consists of several steps. First of them is border detection based on three-channel color and two masks. Then, the authors apply the thinning algorithm to decrease the area of the found boundary. The segmentation algorithm is divided into two parts. In the first part, the image is divided into segments as much as possible. In the second part, the segments union algorithm uses the finding neighbor's ID based on the FSM table and applies a link to ID to create a graph. The results of color images segmentation obtained on the basis of the described algorithm are presented.

Keywords: color image; segmentation; quadtree; edge; border; thinning; color reduction; split; merge; pixel

DOI: 10.14357/08696527200410

References

1. Stojmenovic, M., A. S. Montero, and A. Nayak. 2010. Colour and texture based pyramidal image segmentation. *Conference (International) on Audio, Language and Image Processing Proceedings*. IEEE. 778–786.
2. Correa-Tome, F. E., and R. E. Sanchez-Yanez. 2015. Integral split-and-merge methodology for real-time image segmentation. *J. Electron. Imaging* 24(1):1–11.
3. Volkov, V. N., and A. I. Sorokin. 2018. Klassifikatsiya metodov segmentatsii tsvetnykh izobrazheniy [Classification of methods of segmentation color image]. *Informatsionnye sistemy i tekhnologii* [Information Systems and Technologies] 4(108):32–44.
4. Fu, K. S., and J. K. Mui. 1981. A survey on image segmentation, pattern recognition. *Pattern Recogn.* 13(1):3–16.
5. Ferhat, F. A., O. Kerdjidj, K. Messaoudi, and B. Abdelwahhab. 2013. Implementation of Sobel, Prewitt, Roberts edge detection on FPGA. *World Congress in Computer Science, Computer Engineering, and Applied Computing Proceedings*. 1–4.
6. Lucchese, L., and S. K. Mitray. 2001. Color image segmentation: A state-of-the-art survey. *P. Indian National Science Academy* 67(2):207–221.
7. Verges-Llahi, J. 2005. Color constancy and image segmentation techniques for applications to mobile robotics. Barcelona, Spain: Polytechnic University of Catalonia. PhD Thesis. 23 p.
8. Lorenzo-Navarro, J., and M. Hernández-Tejera. 2007. Image segmentation using a modified split and merge technique. *Cybernet. Syst.* 25(1):137–162.
9. Aizawa, K., K. Motomura, S. Kimura, R. Kadowaki, and J. Fan. 2008. Constant time neighbor finding in quadtrees: An experimental result. *3rd Symposium (International) on Communications, Control and Signal Processing Proceedings*. IEEE. 505–510.
10. Kelkar, D., and S. Gupta. 2008. Improved quadtree method for split merge image segmentation. *1st Conference (International) on Emerging Trends in Engineering and Technology Proceedings*. IEEE. 44–47.
11. Sorokin, A. I., V. N. Volkov, and O. V. Shuldesheva. 2019. Sinteticheskiy algoritm obnaruzheniya konturov ob"ektorov na osnove RGB-koordinat [Synthetic algorithm for detection of object contours on the raster images based on RGB coordinates]. *Informatsionnye sistemy i tekhnologii* [Information Systems and Technologies] 2(112):27–34.
12. Kwon, J. S., J. Woong, and E. K. Kang. 2001. An enhanced thinning algorithm using parallel processing. *Conference (International) on Image Processing Proceedings*. IEEE. 752–755.
13. Laganière, R. 2011. *OpenCV 2 computer vision application programming cookbook*. Packt Publishing. 306 p.
14. Yoder, R., and P. Bloniarz. 2006. A practical algorithm for computing neighbors in quadtrees, octrees, and hyperoctrees. *Conference (International) on Modeling, Simulation & Visualization Methods Proceedings*. CSREA Press. 249–255.

Received September 14, 2020

Contributors

Maniakov Yuri A. (b. 1984)— senior scientist, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; maniakov_yuri@mail.ru

Sorokin Andrey I. (b. 1987)— junior scientist, Orel Branch of the Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 137 Moskovskoe Shosse, Orel 302025, Russian Federation; webdi@mail.ru

ПРОГРАММНЫЙ КОМПЛЕКС ДЛЯ ПОЗИЦИОНИРОВАНИЯ АКУСТИЧЕСКИХ ДОННЫХ СИСТЕМ

*В. А. Смирнов¹, Н. Н. Скворцова², Е. М. Кончеков³, В. А. Ларичев⁴,
Г. А. Максимов⁵*

Аннотация: На основе алгоритмов корреляционного и спектрального анализа разработан программный комплекс для обработки радиофизических сигналов с целью точного позиционирования приемных элементов донных сейсмических кос. Сейсмокосы были испытаны в Геленджикской бухте на Черном море. Представлен результат позиционирования системы, полученный с помощью разработанных алгоритмов по измеренным акустическим дистанциям. Произведено сравнение полученных результатов с контрольными данными измерений эхолота. Показано, что измерение акустических дистанций при использовании разработанного программного комплекса отвечает требованиям для позиционирования систем морской 2D-сейсморазведки.

Ключевые слова: гидроакустика; мониторинг; морская сейсморазведка; твердотельная донная цифровая сейсмокоса; позиционирование; корреляционный анализ; спектральный анализ

DOI: 10.14357/08696527200411

1 Введение

В настоящее время одной из наиболее актуальных задач гидроакустики стала задача сейсмической разведки для целей обнаружения полезных ископаемых на шельфах. Морская сейсморазведка обычно проводится следующими способами: раскладкой донных сейсмических кос, заглублением множества отдельных маяков-ответчиков, буксировкой сейсмических кос у поверхности (глубина $\sim 5\text{--}10$ м) [1]. Получение сейсмических данных при буксировке сейсмических кос у поверхности моря является актуальной как с научной, так и с практической точки зрения задачей [2]. Решение данной задачи связано с разработкой сложных программных комплексов для измерения, обработки и хранения экспериментальных данных [3]. Методы детектирования и обработки таких радиофизических

¹Акустический институт им. академика Н. Н. Андреева; МИРЭА — Российский технологический университет, Viperx15@mail.ru

²Институт общей физики им. А. М. Прохорова Российской академии наук; МИРЭА — Российский технологический университет; Национальный исследовательский ядерный университет «МИФИ», mukudori@mail.ru

³Институт общей физики им. А. М. Прохорова Российской академии наук; Российский национальный исследовательский медицинский университет им. Н. И. Пирогова, eukmek@gmail.com

⁴Акустический институт им. академика Н. Н. Андреева; larido@gmail.com

⁵Акустический институт им. академика Н. Н. Андреева; gamaximov@mail.ru

сигналов необходимы для исследования сигналов и в других областях науки и техники: в акустике, геофизике, исследовании плазменной [4–7] и атмосферной турбулентности [8] и т. д.

Современной и актуальной стала задача обработки лоцирующих акустических сигналов донных и буксируемых акустических антенн (сейсмокос) для получения массива расстояний между датчиками сейсмокосы и источником лоцирующих сигналов для позиционирования системы [9]. В статье будет рассмотрен алгоритм работы созданного программного обеспечения для обработки записей лоцирующих сигналов на приемниках сейсмокосы, выделения времен распространения акустического сигнала по прямому лучу от источника с известными координатами до приемного элемента сейсмокосы для измерения акустических дистанций, позиционирования и вычисления ориентации приемных элементов в пространстве.

2 Позиционирование донной сейсмокосы

В Акустическом институте им. академика Н. Н. Андреева (АО АКИН) разработаны и производятся цифровые твердотельные донные сейсмокосы для инженерной морской сейсморазведки и мониторинга. После заглубления такой косы перед проведением сейсмических измерений требуется определить расположение и ориентацию каждого приемного элемента. Для решения данной задачи в районе раскладки и после заглубления сейсмокосы используется специальное судно-пингер¹, оснащенное спаркером, предназначенным для сейсмических измерений. Для излучения локационных импульсов судно оборудовано высокочастотным пьезокерамическим излучателем, работающим в полосе частот ~ 30 кГц, с помощью которого производится излучение серии высокочастотных широкополосных импульсов. Запись принятых сигналов с приемных элементов донной сейсмокосы синхронизирована с началом излучения судна-пингера.

Работа программы начинается с выбора пользователем директории, содержащей записи сигналов приемных элементов сейсмокосы в формате SEGY 2.0 и записи координат GPS траектории движения судна-пингера с точками, соответствующими моментам излучения лоцирующего сигнала (хлопка). Общая траектория движения судна-пингера и моменты времени, соответствующие записям с сейсмокосы, приведены на рис. 1. На рис. 1 символом «*» отмечен момент излучения лоцирующего импульса. Значком «□» отмечен момент излучения лоцирующего сигнала, для которого в заданной директории была найдена запись принятого сигнала сейсмокосы. Каждому значку «□» соответствует участок записи временных выборок всех приемных элементов сейсмокосы в одном из файлов SEGY.

¹Пингер — это судно, осуществляющее проход по заданной траектории раскладки погружаемых акустических средств (в данном случае сейсмокосы) и излучающее широкополосные мощные лоцирующие импульсы для позиционирования данных средств.

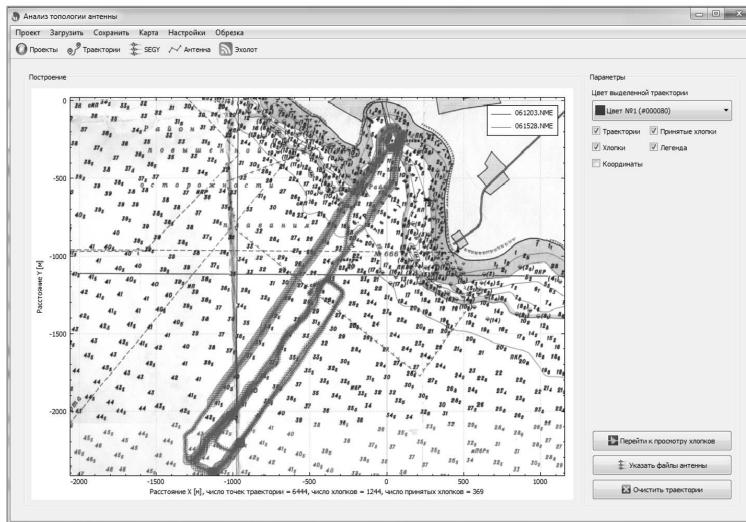


Рис. 1 Траектория движения судна-пингера. По осям X и Y отмечено расстояние рабочего региона в метрах. Общее число точек траектории — 6444. Число моментов излучения лоцирующего импульса — 1244 (значок « $*$ »). Число моментов излучения, для которых проводится анализ записей сигналов — 369 (значок « \square »)

Один приемный элемент сейсмокосы имеет 5 приемных каналов (3 низкочастотных для сейсмических измерений и 2 высокочастотных для позиционирования и поиска ориентации). Частота дискретизации для всех приемных каналов соответствует 2 кГц. На рис. 2 и 3 представлены временные выборки записей всех высокочастотных приемных элементов сейсмокосы при излучении одной серии импульсов, загруженные программой из файлов SEGY. Построение записей сигналов проводится вертикально для визуализации максимального числа каналов на экране пользователя. На рис. 3 представлен небольшой участок для рассмотрения формы лоцирующих импульсов и результатов детектирования. Каждому приемному элементу (гидрофону) соответствуют два приемных канала, что необходимо для дальнейшего вычисления его ориентации в пространстве.

На рис. 2 используются следующие обозначения:

- точками на временных выборках показаны рассчитанные по данной записи времена распространения лоцирующего сигнала по прямому лучу. На рис. 2 приведен пример некорректного детектирования (отброщенное измерение), не прошедшего валидацию (хотя бы по одному критерию). Дистанции, полученные из таких времен, не принимают участия в расчете. Изменяя пороговые значения отсева результатов, можно увеличить число измеренных дистанций, но с уменьшением пороговых значений повышается вероятность попадания некорректно измеренной дистанции в расчет позиционирования;



Рис. 2 Запись сигналов приемных элементов сейсмокосы при излучении одной серии импульсов (приближен участок записи каналов 138–168, только четные каналы)

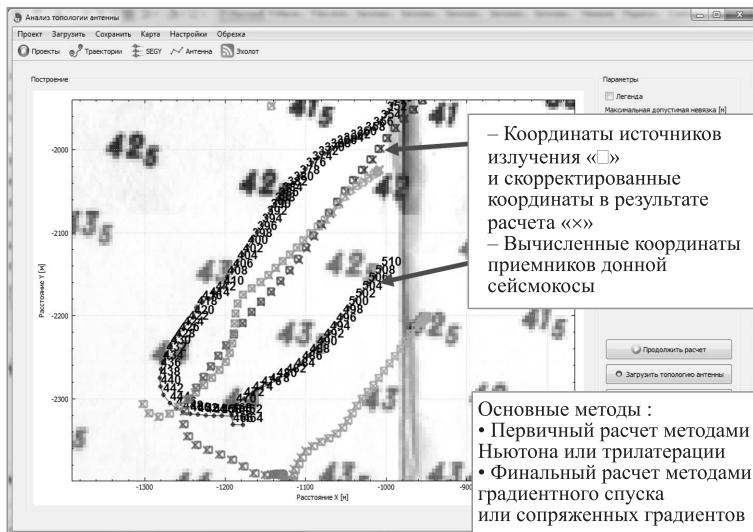


Рис. 3 Финальный результат позиционирования приемных элементов сейсмокосы. По осям X и Y отмечено расстояние рабочего региона в метрах. Черным цветом представлены вычисленные координаты приемных каналов 352–510. Серым цветом представлены измеренные (значок \square) и скорректированные в пределах погрешности координаты GPS- маяка в моменты излучения лоцирующего импульса (значок \times)

- значком « \times » отмечены времена распространения, восстановленные из уже рассчитанного расположения элементов (по всем ~ 1200 записям), они появляются после первого предварительного анализа всех записей и вычисления предварительного результата позиционирования;
- символом « $+$ » показаны начало и конец корреляционных интервалов и значения корреляционных коэффициентов сигналов двух приемных каналов каждого элемента, необходимые для поиска ориентации приемного элемента в пространстве, цифрой указан уровень корреляции сигналов.

Расчет времен прямого распространения лоцирующего сигнала проводится следующими методами численного анализа временных выборок.

1. Вычисление отношения сигнал/шум на всем интервале сигнала для выбора рабочего участка.
2. Корреляция огибающей принятого сигнала и огибающей излученной последовательности лоцирующих импульсов. Значения ниже порогового уровня корреляции отбрасываются.
3. Медианная фильтрация¹ принятого сигнала с заданным окном для коррекции времен распространения сигналов и борьбы с замираниями.

Решение задачи позиционирования проходит в два этапа.

1. Вычисление предварительного результата с помощью метода Ньютона или трилатерации.
2. Вычисление конечного результата с помощью итерационных методов: градиентного спуска² или метода сопряженных градиентов³.

После запуска расчета программа автоматически анализирует все сопоставленные моменты излучения с записями сигналов, проводит измерение акустических дистанций и представляет результат позиционирования. На рис. 3 представлен конечный результат позиционирования приемных элементов донной сейсмокосы.

¹Медианный фильтр — один из видов цифровых фильтров, широко используемый в цифровой обработке сигналов и изображений для уменьшения уровня шума. Медианный фильтр является нелинейным фильтром с конечной импульсной характеристикой. Значения отсчетов внутри окна фильтра сортируются в порядке возрастания (убывания); и значение, находящееся в середине упорядоченного списка, поступает на выход фильтра. В случае четного числа отсчетов в окне выходное значение фильтра равно среднему значению двух отсчетов в середине упорядоченного списка. Окно перемещается вдоль фильтруемого сигнала, и вычисления повторяются.

²Градиентный спуск — метод нахождения локального экстремума функции с помощью движения вдоль градиента.

³Метод сопряженных градиентов — метод нахождения локального экстремума функции на основе информации о ее значениях и ее градиенте.

3 Ориентация приемных элементов донной сейсмокосы

Задача поиска ориентации решается путем построения математической модели распространения акустического сигнала с известными частотными характеристиками от источника к двум гидрофонам приемника, находящимся на известном расстоянии друг от друга (14–17 см). Затем осуществляется перебор значений входных параметров модели:

- горизонтальное расположение приемника (перебор с точностью $0,5^\circ$);
- вертикальное расположение приемника (перебор с точностью $0,5^\circ$);
- дистанция между приемниками одного элемента (перебор с точностью 0,5 мм);
- частота несущей излучаемого сигнала (~ 30 кГц).

Цель перебора для данной задачи заключается в поиске параметров модельной функции, для которой корреляционные коэффициенты сигналов приемных каналов одного элемента из разных источников совпадают в пределах допустимой погрешности с экспериментальными значениями, определяемыми в ходе анализа сигналов приемных каналов после решения задачи позиционирования.

Расположение элемента, при котором модельные коэффициенты совпадают с экспериментальными в пределах допустимой погрешности, считается правильным ответом. Если такое расположение не было найдено при переборе всех доступных параметров моделирования, ориентация остается не определенной и отмечается серым цветом.

Результат сравнения полученной математической модели, учитывающей расположение приемного элемента, с исходными (экспериментальными) коэффициентами представлен на рис. 4. Коррелируемые сигналы представлены справа.

На рис. 5 представлен результат вычисления ориентации приемных элементов сейсмокосы в пространстве (серым цветом). Серое значение соответствует дистанции между приемными каналами приемника в метрах (на этапе моделирования это один из параметров, который может изменяться в небольших пределах).

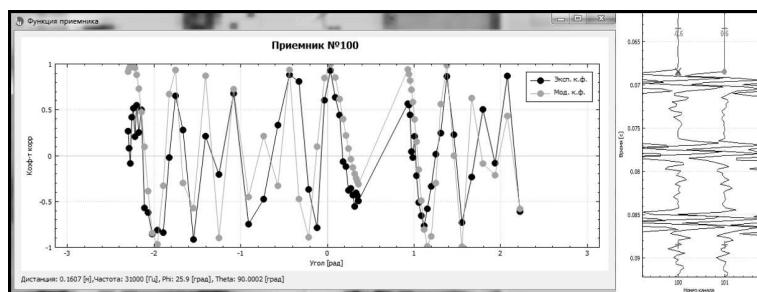


Рис. 4 Сравнение функции распределения взаимно корреляционных коэффициентов приемных каналов одного элемента (модель и эксперимент)

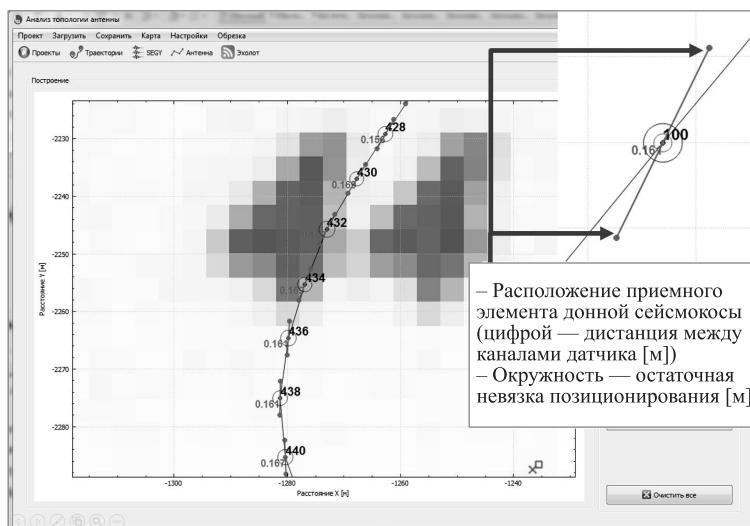


Рис. 5 Ориентация элементов донной сейсмокосы

Таким образом, можно применить разработанные алгоритмы спектрально-корреляционной обработки радиофизических сигналов для задачи детектирования времен прямого распространения акустических лоцирующих сигналов для решения задачи позиционирования и ориентации приемных элементов донной сейсмокосы.

Полученные результаты позиционирования сравнивались с данными глубин эхолота, установленного на судне-пингере. Конечная ошибка позиционирования не может быть меньше точности GPS-маяка, установленного на судне (1–3 м), однако при большом числе точных измерений дистанций (более 50 для каждого приемного элемента) статистически ошибка позиционирования может быть уменьшена до 0,5 м и менее. При расчете позиционирования существует возможность зафиксировать глубину, замещая ее данными эхолота, что может уменьшить ошибку позиционирования до 0,1 м.

4 Заключение

Реализовано бортовое кроссплатформенное бортовое программное обеспечение (протестировано на Windows XP, Windows 7, Windows 10, Ubuntu 18, Debian 9 и 10) для визуализации и сопоставления данных траекторий движения судна-пингера, обработки и анализа данных акустических сигналов донных сейсмокос, позиционирования и поиска ориентации приемных элементов донных сейсмокос. Программное обеспечение создавалось средствами C++ 17 с по-

мошью открытого пакета для разработки QT Creator 5. Реализованы алгоритмы для выделения времен прямого распространения акустических сигналов от источника с известными координатами (пингера) до приемных элементов донной сейсмокосы, реализован ряд алгоритмов для борьбы с замираниями и помехами. Проведено сравнение полученных результатов позиционирования с данными эхолота. Полученная точность позиционирования достаточна для задач морской 2D-сейсморазведки.

Литература

1. Максимов Г. А., Денисов Д. Н., Ларичев В. А., Лесонен Д. Н., Григорьев А. Г., Корольков З. А. Цифровая твердотельная буксируемая сейсмокоса малого диаметра для морской инженерной сейсморазведки // Учен. зап. физ. фак-та Моск. ун-та, 2017. № 5. Ст. № 1750126.
2. Амтилов Ю. П. Освоение шельфа Арктики и Дальнего Востока. Проблемы и перспективы // Offshore Russia, 2014. № 4. С. 8–15.
3. Малахов Д. В., Смирнов В. А., Скворцова Н. Н., Соколов А. С., Степахин В. Д., Борзосеков В. Д., Харчев Н. К. Радиофизический программно-аппаратный комплекс сбора и обработки данных на стеллараторе Л-2М // Инженерная физика, 2015. № 8. С. 38–44.
4. Skvortsova N. N., Akulina D. K., Batanov G. M., Kharchev N. K., Kolik L. V., Kovrizhnykh L. M., Letunov A. A., Logvinenko V. P., Malakhov D. V., Petrov A. E., Pshenichnikov A. A., Sarkсян K. A., Voronov G. S. Effect of ECRH regime on characteristics of short-wave turbulence in plasma of the L-2M // Plasma Phys. Contr. F., 2010. Vol. 52. Art. No. 055008.
5. Батанов Г. М., Борзосеков В. Д., Колик Л. В., Малахов Д. В., Петров А. Е., Пшеничников А. А., Сарксян К. А., Скворцова Н. Н., Харчев Н. К. Длинноволновая турбулентность в плазме стелларатора Л-2М при электронно-циклонном нагреве // Вопросы атомной науки и техники. Сер.: Термоядерный синтез, 2011. № 2. С. 70–75.
6. Батанов Г. М., Борзосеков В. Д., Коврижных Л. М., Колик Л. В., Кончеков Е. М., Малахов Д. В., Петров А. Е., Сарксян К. А., Скворцова Н. Н., Степахин В. Д., Харчев Н. К. Рассеяние назад излучения гиротрона при ЭЦ нагреве плазмы на стеллараторе Л-2М и коротковолновая турбулентность // Физика плазмы, 2013. Т. 39. № 6. С. 511–522.
7. Скворцова Н. Н., Степахин В. Д., Малахов Д. В., Сорокин А. А., Батанов Г. М., Борзосеков В. Д., Глявин М. Ю., Колик Л. В., Кончеков Е. М., Летунов А. А., Петров А. Е., Рябикова И. Г., Сарксян К. А., Соколов А. С., Смирнов В. А., Харчев Н. К. Создание рельефа на молибденовых пластинах в разрядах, инициируемых излучением гиротрона в порошках металл–диэлектрик // Известия вузов. Радиофизика, 2015. Т. 58. № 9. С. 779–788.
8. Маслов С. А., Харчевский А. А., Смирнов В. А. Применение вейвлета Хаара для анализа плазменных и атмосферных флуктуаций // Ядерная физика и инжиниринг, 2016. Т. 7. № 5. С. 448–452.

9. Смирнов В. А., Скворцова Н. Н., Максимов Г. А., Ларичев В. А., Смагин Д. А., Лекомцев В. М. Алгоритмы обработки радиофизических сигналов для систем дистанционного мониторинга в гидросфере // Прикладная физика, 2019. № 5. С. 85–92.

Поступила в редакцию 20.02.20

SOFTWARE PACKAGE FOR POSITIONING OF ACOUSTIC BOTTOM SYSTEMS

**V.A. Smirnov^{1,2}, N.N. Skvortsova^{2,3,4}, E.M. Konchekov^{3,5}, V.A. Larichev¹,
and G.A. Maximov¹**

¹N.N. Andreev Acoustic Institute, 4 Shverniaka Str., Moscow 117036, Russian Federation

²MIREA — Russian Technological University, 78 Vernadskogo Ave., Moscow 119454, Russian Federation

³Prokhorov General Physics Institute of the Russian Academy of Sciences, 38 Vavilov Str., Moscow 119991, Russian Federation

⁴National Research Nuclear University “MEPhI,” 31 Kashirskoe Shosse, Moscow 115409, Russian Federation

⁵Pirogov Russian National Research Medical University, 1 Ostrovitianov Str., Moscow 117997, Russian Federation

Abstract: The problem of accurate positioning of the receiving elements of the bottom seismic streamers is considered. A new software package for processing, based on the algorithms of correlation and spectral analysis, of radio-physical signals is developed. Seismic streamers were tested in the Gelendzhik bay on the Black Sea. The system positioning results, obtained using the developed algorithms according to the measured acoustic distances, are presented and compared with the echo sounder measurement data. It is shown that measurements of acoustic distances using the developed software meet the requirements for positioning of two-dimensional marine seismic systems.

Keywords: hydroacoustics; monitoring; marine seismic; solid-state bottom digital seismic streamer; positioning; correlation analysis; spectral analysis

DOI: 10.14357/08696527200411

References

1. Maximov, G. A., D. M. Denisov, V. A. Larichev, D. N. Lesonen, A. G. Grigor'ev, and Z. A. Korolkov. 2017. Tsifrovaya tverdotel'naya buksiruyemaya seysmokosa malogo diametra dlya morskoy inzhenernoy seysmorazvedki [Digital solid seismic streamer of a small diameter for the marine engineering seismic exploration]. *Uchen. zap. fiz. fak-ta Mosk. un-ta* [Moscow University Physics Bull.] 5:1750126.

2. Ampilov, Yu. P. 2014. Osvoenie shel'fa Arktiki i Dal'nego Vostoka. Problemy i perspektivy [Developing the arctic shelf and far eastern shelf. Problems and perspectives]. *Offshore Russia* 4:8–15.
3. Malakhov, D. V., V. A. Smirnov, N. N. Skvortsova, A. S. Sokolov, V. D. Stepakhan, V. D. Borzosekov, and N. K. Harchev. 2015. Radiofizicheskiy apparatnyy kompleks sbora i obrabotki dannykh na stellaratore L-2M [Radiophysical hardware and software system for acquisition and data processing on stellarator L-2M]. *Inzhenernaya fizika* [Engineering Physics] 8:38–44.
4. Skvortsova, N. N., D. K. Akulina, G. M. Batanov, N. K. Kharchev, L. V. Kolik, L. M. Kovrzhnykh, A. A. Letunov, V. P. Logvinenko, D. V. Malakhov, A. E. Petrov, A. A. Pshenichnikov, K. A. Sarkisyan, and G. S. Voronov. 2010. Effect of ECRH regime on characteristics of short-wave turbulence in plasma of the L-2M stellarator. *Plasma Phys. Control. F.* 52:055008.
5. Batanov, G. M., V. D. Borzosekov, L. V. Kolik, D. V. Malakhov, A. E. Petrov, A. A. Pshenichnikov, K. A. Sarkisyan, N. N. Skvortsova, and N. K. Kharchev. 2011. Dlinnovolnovaya turbulentnost' v plazme stellaratora L-2M pri elektronnotsiklotronnom nagreve [Long-wavelength turbulence during ECR heating of the plasma in L-2M stellarator]. *Voprosy atomnoy nauki i tekhniki. Ser. Termoyadernyy sintez* [Problems of Atomic Science and Technology. Thermonuclear Fusion Ser.] 2:70–75.
6. Batanov, G. M., V. D. Borzosekov, L. M. Kovrzhnykh, L. V. Kolik, E. M. Konchekov, D. V. Malakhov, A. E. Petrov, K. A. Sarkisyan, N. N. Skvortsova, V. D. Stepakhan, and N. K. Kharchev. 2013. Backscattering of gyrotron radiation and short-wavelength turbulence during electron cyclotron resonance plasma heating in the L-2M stellarator. *Plasma Phys. Rep.* 39(6):444–455.
7. Skvortsova, N. N., V. D. Stepakhan, D. V. Malakhov, A. A. Sorokin, G. M. Batanov, V. D. Borzosekov, M. Yu. Glyavin, L. V. Kolik, E. M. Konchekov, A. A. Letunov, A. E. Petrov, I. G. Ryabikina, K. A. Sarkisyan, A. S. Sokolov, V. A. Smirnov, and N. K. Kharchev. 2015. Relief creation on molybdenum plates in discharges initiated by gyrotron radiation in metal–dielectric powder mixtures. *Radiophys. Quantum El.* 58(9):701–709.
8. Maslov, S. A., A. A. Kharchevsky, and V. A. Smirnov. 2017. Application of the Haar wavelet to the analysis of plasma and atmospheric fluctuations. *Phys. Atom. Nucl.* 80(11): 1692–1696.
9. Smirnov, V. A., N. N. Skvortsova, G. A. Maksimov, V. A. Larichev, D. A. Smagin, and V. M. Lekomcev. 2019. Algoritmy obrabotki radiofizicheskikh signalov dlya sistem distantsionnogo monitoringa v gidrosfere [Data processing algorithms of radiophysical signals for remote monitoring systems in hydrosphere]. *Prikladnaya fizika* [Applied Physics] 5:85–92.

Received February 20, 2020

Contributors

Smirnov Vitaliy A. (b. 1990) — PhD student, Department of Modeling Radio-Physical Processes, MIREA — Russian Technological University, 78 Vernadskogo Ave., Moscow 119454, Russian Federation; leading engineer, Department of

Acoustic Transducers, N. N. Andreev Acoustic Institute, 4 Shvernika Str., Moscow 117036, Russian Federation; Viperx15@mail.ru

Skvortsova Nina N. (b. 1956) — Doctor of Science in physics and mathematics, professor, Department of Modeling Radio-Physical Processes, MIREA — Russian Technological University, 78 Vernadskogo Ave., Moscow 119454, Russian Federation; professor, National Research Nuclear University “MEPhI,” 31 Kashirskoe Shosse, Moscow 115409, Russian Federation; leading scientist, Prokhorov General Physics Institute of the Russian Academy of Sciences, 38 Vavilov Str., Moscow 119991, Russian Federation; mukudori@mail.ru

Konchekov Evgeny M. (b. 1988) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Physics, Pirogov Russian National Research Medical University, 1 Ostrovitianov Str., Moscow 117997, Russian Federation; senior scientist, Prokhorov General Physics Institute of the Russian Academy of Sciences, 38 Vavilov Str., Moscow 119991, Russian Federation; eukmek@gmail.com

Larichev Vladimir A. (b. 1972) — Candidate of Science (PhD) in physics and mathematics, scientist, N. N. Andreev Acoustic Institute, 4 Shvernika Str., Moscow 117036, Russian Federation; larido@gmail.com

Maksimov German A. (b. 1963) — Doctor of Science in physics and mathematics, head of department, N. N. Andreev Acoustic Institute, 4 Shvernika Str., Moscow 117036, Russian Federation; gamaximov@mail.ru

МАШИННЫЙ ПЕРЕВОД: ИНДИКАТОРНАЯ ОЦЕНКА РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ

А. Ю. Егорова¹, И. М. Зацман², М. Г. Кружков³, В. А. Нуриев⁴

Аннотация: Представлены данные, полученные в ходе наблюдения за обучением системы нейронного машинного перевода (НМП). Проведена количественная оценка работы системы НМП с помощью индикаторов. В качестве экспериментального материала были использованы 250 русскоязычных текстовых фрагментов, для каждого из которых ежемесячно в течение одного года фиксировался перевод на французский язык, выполненный с помощью системы НМП Google. Фиксация переводов была реализована посредством их аннотирования в надкорпусной базе данных (НБД), в результате чего была сформирована серия из 12 аннотаций для каждого из 250 текстовых фрагментов. Аннотирование переводов позволило не только зафиксировать допущенные в переводе ошибки в случае их наличия, но и определить категорию нестабильности НМП, указывающую на изменения качества перевода или на их отсутствие. Цель статьи — представить разработанный индикаторный подход и пример его применения для оценки результатов обучения искусственной нейронной сети (ИНС).

Ключевые слова: нейронный машинный перевод; нестабильность машинного перевода; индикаторная оценка; лингвистическое аннотирование; виды нестабильности

DOI: 10.14357/08696527200412

1 Введение

Компьютерные системы после обучения ИНС следуют правилам решения задачи, которые определяются данными в процессе ее обучения (data-driven), а не создаются заранее человеком. Алгоритмы, определяемые данными, не соответствуют правилам, которым следуют люди при решении этой же задачи. Согласно Фенстермахеру, компьютерные системы могут выполнять ту же задачу, используя альтернативные правила и другие формы представления знания, при-

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ann.shurova@gmail.com

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, izatsman@yandex.ru

³Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, magnit75@yandex.ru

⁴Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, nurieff.v@gmail.com

меняется при ее решении [1]. Например, машинный алгоритм распознавания лиц, сформированный в процессе обучения компьютерной системы, не совпадает с алгоритмом, разработанным человеком [2].

Основная причина отличий машинных и разработанных человеком алгоритмов заключается в следующем. Искусственная нейронная сеть состоит из слоев искусственных нейронов, соединенных между собой взвешенными связями. Искусственные нейронные сети не программируются обычным способом согласно алгоритму, разработанному человеком. Они обучаются на большом числе учебных примеров с их решениями (ответами) и настраивают весовые коэффициенты (веса) связей так, чтобы ИНС для каждого обучающего примера давала правильный ответ. При этом объем матрицы весов связей может составлять несколько гигабайт. После обучения ИНС используется для решения задач, не входящих в обучающие примеры. Что произойдет, если ИНС даст неправильный результат при их решении и возникнет необходимость узнать, почему произошла ошибка, является ли она постоянной или нестабильной? В традиционной программе человек ищет сегмент кода, ответственный за ошибку, анализирует ее и исправляет. В ИНС нет алгоритмических шагов, и можно увидеть только гигабайтную матрицу весов. Эта матрица весов непрозрачна — с ее помощью нельзя выяснить, как веса соотносятся с ошибкой. Вот почему анализ и оценка результатов, получаемых при обучении ИНС, становятся в ряд самых актуальных областей исследований в информатике [3] и компьютерной лингвистике.

В статье обучение ИНС демонстрируется на материале русско-французских аннотаций (см. примеры с коннекторами¹ *как (расстояние)*, *так и* в табл. 4 и *не только (расстояние)Ø* в табл. 5). Аннотации были созданы во время эксперимента с помощью НБД коннекторов [5–10]. Каждая аннотация включает текстовый фрагмент с коннектором на русском языке, перевод фрагмента на французский, выполненный с помощью системы НМП Google, и структурированное описание ошибок в случае их наличия.

Для эксперимента были отобраны 250 текстовых фрагментов на русском языке, каждый из которых содержит двухкомпонентный коннектор (см., например, *как (расстояние)*, *так и* в табл. 4). В эксперименте использовались тексты на русском языке и их французские переводы, полученные из французского подкорпуса Национального корпуса русского языка (НКРЯ) [11].

На протяжении года, с марта 2019 г. по февраль 2020 г., для отобранных фрагментов в НБД коннекторов ежемесячно формировались и регистрировались 250 аннотаций. В результате был получен массив экспериментальных данных, в котором для каждого из 250 текстовых фрагментов была зафиксирована серия из 12 версий машинного перевода. Таким образом, всего в НБД было создано 3000 аннотаций.

¹ Коннектор — языковая единица, функция которой состоит в выражении логико-семантического отношения, существующего между соединенными с ее помощью частями текстового фрагмента [4, с. 17].

Для обработки сформированных аннотаций применена оценка результатов обучения ИНС с использованием системы количественных показателей (индикаторов), значения которых вычисляются на основе мониторинга работы ИНС (далее — индикаторная оценка результатов обучения). Система индикаторов и методика мониторинга были разработаны в три этапа. На первом этапе, описанном в [12], была проведена категоризация нестабильности НМП на основе анализа ошибок перевода, которые возникали или исчезали в сериях машинного перевода при обучении ИНС. В результате были выделены 6 категорий — 5 из них характеризуют нестабильность качества НМП:

- повышение качества НМП;
- снижение качества НМП;
- колебание качества НМП;
- изменение набора ошибок в НМП без динамики его качества;
- изменение НМП без динамики его качества.

Шестая категория объединяет переводы, которые не менялись в пределах серии (подробное описание категорий дано в [12]).

На втором этапе была разработана и описана методика темпоральной оценки нестабильности НМП [13]. Настоящая статья посвящена третьему этапу исследования, цель которого — оценка результатов обучения ИНС на основе мониторинга ее работы с применением разработанной системы индикаторов (табл. 1).

2 Нестабильность нейронного машинного перевода и виды ошибок перевода

В результате эксперимента для каждой из 6 категорий были получены числовые данные для случаев, когда ошибки в машинном переводе отсутствуют (см. строку 1 табл. 1), а также данные о частотности 19 видов ошибок машинного перевода¹ (см. строки 2–20 табл. 1). Частотность вычислялась от общего числа ошибок (2836), зарегистрированных в 1800 аннотациях, так как одна аннотация могла содержать описание нескольких ошибок.

На момент написания настоящей статьи была проведена рубрикация по видам ошибок и категориям, характеризующим нестабильность НМП, для 60% сформированных в НБД аннотаций (1800 из 3000). На основании анализа полученных данных можно выявить ряд закономерностей.

По сумме ошибок во всех 6 категориях, характеризующих нестабильность НМП, самой частотной стала ошибка ErrorSemant — лексическая ошибка (38,26% от 2836 ошибок), искажающая смысл оригинала в переводном фрагменте (см. также строку 11 в табл. 4 и строки 6 и 11 в табл. 5). В НБД было зарегистрировано 1085 аннотаций с такой ошибкой из 1800.

¹ Подробнее о классификации ошибок см. в [12, 13].

Таблица 1 Индикаторы нестабильности НМП, ошибок и случаев их отсутствия

№	Код ошибки	Повы- шение каче- ства НМП	Сни- жение каче- ства НМП	Коле- бание каче- ства НМП	Измене- ние набора ошибок в НМП без дина- мики его качества	Изме- нение НМП без дина- мики его качества	НМП без изме- нений	Сумма по 6 кате- гориям, характери- зующим неустабиль- ность	Значе- ния инди- каторов ошибок, %
		1	2	3	4	5	6		
1	NoError	73	66	82	0	156	48	425	14,99
2	ErrorSemant	366	67	362	217	61	12	1085	38,26
3	ErrorSyntax	108	25	94	66	12	0	305	10,75
4	ErrorSyntaxPostCNT	82	19	60	46	24	0	231	8,15
5	ErrorTotalCNT	64	3	19	62	24	0	172	6,06
6	ErrorCNT	46	0	33	55	0	0	134	4,72
7	Pleonasm	28	12	15	23	11	0	89	3,14
8	ErrorMorphPostCNT	16	0	17	27	0	0	60	2,12
9	ErrorMorphCNT	16	0	2	20	12	0	50	1,76
10	Lacuna	13	0	16	36	0	0	65	2,29
11	ErrorPart2CNT	26	3	6	16	12	0	63	2,22
12	ErrorPunct	9	3	21	10	0	0	43	1,52
13	ErrorPart1CNT	15	0	16	13	12	0	56	1,97
14	ErrorMorph	5	3	14	13	0	0	35	1,23
15	AgramTotal	5	0	0	7	0	0	12	0,42
16	TrPart1CNT	0	0	5	0	0	0	5	0,18
17	Latin	5	0	0	0	0	0	5	0,18
18	TrPart2CNT	0	0	1	0	0	0	1	0,04
19	Cyrillic	0	0	0	0	0	0	0	0,00
20	ErrorOrthCNT	0	0	0	0	0	0	0	0,00
Сумма по видам ошибок		877	201	763	611	324	60	2836	
Значения индикаторов для каждой категории, %		30,92	7,09	26,90	21,54	11,42	2,12		100,00

Далее по уменьшению частотности (пропуская код NoError, характеризующий допустимые варианты перевода) — синтаксические ошибки во фрагменте, не вводимом коннектором (ErrorSyntax — 10,75%) и вводимом коннектором (ErrorSyntaxPostCNT — 8,15%). Пример перевода с ошибкой ErrorSyntaxPostCNT указан в строке 1 табл. 4. Морфологических ошибок (ErrorMorphPostCNT для фрагмента с коннектором и ErrorMorph для фрагмента, не вводимого коннектором), напротив, значительно меньше — 2,12% и 1,76% соответственно. Это свидетельствует о том, что система НМП в основном хорошо справляется с грамматикой на уровне морфологии и трудности возникают чаще на уровне синтаксиса. Однако устойчивое снижение числа морфологических ошибок в период проведения эксперимента не наблюдалось (табл. 2).

Таблица 2 Динамика ошибок по месяцам

Код ошибки	Месяц											
	1	2	3	4	5	6	7	8	9	10	11	12
AgramTotal	3	2	3	2	2	0	0	0	0	0	0	0
ErrorMorph	3	2	3	2	2	2	3	3	5	4	4	4
ErrorMorphPostCNT	5	5	7	6	7	4	4	4	5	5	4	4
ErrorSyntax	28	27	28	26	27	23	22	21	24	28	25	26
ErrorSyntaxPostCNT	27	24	22	25	22	15	16	16	16	16	16	16
ErrorSemant	85	84	82	83	83	93	94	95	96	97	95	98
ErrorPunct	4	5	4	4	4	3	3	3	3	2	4	4
Cyrillic	0	0	0	0	0	0	0	0	0	0	0	0
Latin	1	1	1	1	1	0	0	0	0	0	0	0
Lacuna	8	5	8	5	4	6	6	6	6	5	3	3
Pleonasm	6	7	9	10	8	7	8	8	7	6	6	7
TrPart1CNT	0	0	0	0	0	2	1	1	1	0	0	0
TrPart2CNT	0	0	0	1	0	0	0	0	0	1	1	1
ErrorPart1CNT	5	5	6	6	6	2	2	2	3	4	5	5
ErrorPart2CNT	10	10	8	8	8	2	2	2	2	2	2	2
ErrorTotalCNT	20	21	22	22	23	7	7	8	9	8	9	9
ErrorMorphCNT	6	5	5	6	5	3	3	2	2	3	4	6
ErrorOrthCNT	0	0	0	0	0	0	0	0	0	0	0	0
ErrorCNT	13	13	14	14	15	9	9	9	11	9	9	9
NoError	32	34	36	37	35	39	39	38	36	32	34	33

Среди ошибок, касающихся непосредственно перевода коннекторов, самыми распространенными стали ErrorTotalCNT (система НМП переводит коннектор несуществующей языковой единицей, пример см. в строке 1 табл. 4) и ErrorCNT (система определяет, что в переводе нужен коннектор, но ошибается в его выборе, см. строку 6 табл. 4).

Хотя система НМП нередко допускает ошибки при переводе коннектора на французский язык, по данным из НБД коннекторов, не вошедшим в табл. 1, было зарегистрировано относительно мало случаев, когда коннектор не переведен вовсе, всего 34 аннотации из 1800. Объем настоящей статьи не позволяет описать это наблюдение подробнее, отметим только, что перевод русскоязычного коннектора нулевым эквивалентом не всегда ошибочен.

На данном этапе нулевыми остались ячейки для ошибок Cyrillic (слова кириллицей в переводе) и ErrorOrthCNT (орфографическая ошибка в форме коннектора). Это может указывать на тенденцию к повышению качества НМП, поскольку в предыдущих исследованиях [14] эти значения не были нулевыми.

Кроме того, рассматривая данные для первых пяти категорий, которые характеризуют нестабильность НМП, можно отметить следующее: первой категории «Повышение качества НМП» соответствует наибольшее число ошибок (30,92%),

а второй категории «Снижение качества НМП» — наименьшее число ошибок (7,09%), что заслуживает отдельного лингвистического анализа, выходящего за рамки этой статьи.

3 Динамика ошибок на протяжении эксперимента

Использование НБД при изучении нестабильности НМП позволило не только получить значения индикаторов по категориям, характеризующим нестабильность НМП, и по ошибкам НМП, но и проследить динамику ошибок каждого вида на протяжении всего эксперимента (см. табл. 2).

Данные табл. 2 свидетельствуют о том, что у разных видов ошибок разная динамика. Основная тенденция (строки для ошибок ErrorSyntaxPostCNT, Lacuna, ErrorPart2CNT, ErrorTotalCNT и ErrorCNT¹) — уменьшение числа ошибок со временем. Однако есть ошибки (ErrorSyntax, Pleonasm² и ErrorMorphPostCNT, ErrorMorph³), число которых колеблется. Также есть виды ошибок, число которых, напротив, увеличивается (например, ErrorSemant⁴ на 15,29%).

4 Распределение аннотаций по категориям, характеризующим нестабильность нейронного машинного перевода

Тенденции, продемонстрированные в табл. 2, коррелируют с данными табл. 3, в которой представлено распределение аннотаций по категориям, характеризующим нестабильность НМП.

Так, наибольшее значение индикатора (30,6% аннотаций) у 1-й категории, характеризующей повышение качества НМП (пример см. в табл. 4). За ней по частотности следуют рубрики, сигнализирующие об изменении НМП без однозначной динамики его качества к повышению или понижению. Это колебание качества НМП (3-я категория), изменение набора ошибок в НМП без динамики его качества (4-я категория) и изменение НМП без динамики его качества (5-я категория). Наименьшее число аннотаций приходится на 2-ю категорию (снижение качества НМП) — 9,3% аннотаций (пример см. в табл. 5). Стабильными в течение эксперимента остались лишь 3,3% переводов (НМП без изменений).

В табл. 4 представлена серия переводов следующего текстового фрагмента: «Как воровали, так и будут воровать» [Светлана Алексиевич. Время секунд хэнд

¹ErrorSyntaxPostCNT — синтаксическая ошибка во фрагменте текста, вводимом коннектором; Lacuna — пропуск фрагмента текста; ErrorPart2CNT — вторая часть неоднословного коннектора переведена ошибочно; ErrorTotalCNT — коннектор переведен несуществующей языковой единицей; ErrorCNT — семантическая ошибка в выборе коннектора.

²ErrorSyntax — синтаксическая ошибка во фрагменте текста, не вводимом коннектором; Pleonasm — избыточный перевод.

³ErrorMorphPostCNT — морфологическая ошибка во фрагменте текста, вводимом коннектором; ErrorMorph — морфологическая ошибка во фрагменте текста, не вводимом коннектором.

⁴ErrorSemant — лексическая ошибка, искажение смысла переводимого фрагмента.

Таблица 3 Распределение аннотаций по категориям, характеризующим нестабильность НМП

№	Название рубрики, характеризующей нестабильность НМП	Число аннотаций	Доля от общего числа аннотаций (1800)
1	Повышение качества НМП	552	30,6%
2	Снижение качества НМП	168	9,3%
3	Колебание качества НМП	492	27,3%
4	Изменение набора ошибок в НМП без динамики его качества	300	16,6%
5	Изменение НМП без динамики его качества	228	12,6%
6	НМП без изменений	60	3,3%

(ч. 2) (2013)]¹. По итогам проведенного лингвистического анализа, методика которого подробно описана в [13], для этой серии переводов была простоявленна категория «Повышение качества НМП».

За год эксперимента по изучению нестабильности НМП, которая, по всей видимости, вызвана обучением ИНС, машинный перевод исходного текстового фрагмента менялся три раза (повторяющиеся значения таблицы заменены символом -// -). В первой версии перевода от 03.03.19 допущено три ошибки. Это ErrorSyntaxPostCNT (отсутствует синтаксическая связь между клаузами *comment voler* и *alors ils voleront*), а также две ошибки в переводе коннектора — ErrorTotalCNT (русский коннектор *как||так* и переведен несуществующей во французском языке единицей *comment||alors*) и ErrorMorph (часть коннектора ошибочно заменена языковой единицей *comment*, не являющейся коннектором).

В 6-м переводе от 03.08.19 допущена всего одна ошибка — это ErrorCNT (выбран семантически неверный коннектор *comme*). Подобные ошибки зачастую ведут кискажению смысла всего переводимого фрагмента: при использовании коннектора *comme* фраза приобретает смысл «Потому как они воровали, они будут воровать».

В 11-м переводе от 02.01.20 допущена ErrorSemant: при переводе коннектора *как||так* и используется грамматическая структура *gérondif présent*, и фраза приобретает смысл «Воруя, они будут воровать».

Несмотря на то что в последнем, 12-м, переводе сохраняется ErrorSemant, в целом в этой серии переводов прослеживается тенденция к повышению качества машинного перевода, так как число ошибок со временем уменьшается и переводной текст становится более понятным.

¹Профессиональный перевод этого предложения на французский язык: «Les gens continueront à voler comme avant» [Sophie Benech, 2013].

Таблица 4 Серия аннотаций с повышением качества НМП

№ п/п	№ в НБД	Контекст РР в оригинале ¹	РР в ори- гинале и ее при- знаки	Контекст РР в переводе ²	РР в переводе и ее признаки	При- знаки анно- тации	Дата и время записи в НБД
1	28312	Как воровали, так и будут воровать.	как так и (аналогия) (CNT p CNT q) (CNT) (Дистант)	Comment voler, alors ils voleront.	comment alors (ErrorSyntaxPostCNT) (ErrorTotalCNT) (ErrorMorphCNT)	MT	03.03.19 17:20
2	28670		- // -	- // -	- // -	- // -	01.04.19 15:39
3	29324		- // -	- // -	- // -	- // -	05.05.19 14:38
4	29913		- // -	- // -	- // -	- // -	05.06.19 14:26
5	30174		- // -	- // -	- // -	- // -	02.07.19 13:45
6	30461		Comme ils ont volé, ils vont voler.	comme (ErrorCNT)	Cngrn MT	03.08.19 23:37	
7	30707		- // -	- // -	- // -	- // -	03.09.19 15:04
8	31022		- // -	- // -	- // -	- // -	02.10.19 22:23
9	31545		- // -	- // -	- // -	- // -	04.11.19 15:27
10	32078		- // -	- // -	- // -	- // -	02.12.19 15:36
11	32485		En volant, ils voleront.	x_gér_prés (ErrorSemant)	Dvrg MT	02.01.20 20:02	
12	33171		- // -	x_gér_prés (ErrorSemant) (Повышение качества НМП)	- // -	- // -	03.02.20 15:50

¹Речевая реализация (РР) в оригинале — форма коннектора в конкретном высказывании [15].²Речевая реализация в переводе — функционально эквивалентный фрагмент перевода, передающий смысловое наполнение РР в оригинале. Понятие «функционально эквивалентный фрагмент» предложено в [16].**Примечания.**

1. Рубрика (подробнее о рубриках см. в [7]) «аналогия» говорит о том, что коннектор *как (расстояние), так и* выражает логико-семантическое отношение аналогии.
2. Рубрика «CNT p CNT q» говорит о том, что элементы двухкомпонентного коннектора находятся в каждом из соединяемых компонентов текста *p* («воровали») и *q* («будут воровать»).
3. Рубрика «CNT» говорит о том, что аннотация сформирована для всего коннектора, а не для отдельно составляющих его блоков или элементов.
4. Рубрика «Дистант» говорит о том, что составные части коннектора «как» и «так и» разделены текстом «воровали».
5. Рубрика «МТ» проставляется автоматически при использовании машинного переводчика.
6. Рубрика «Cngrn» говорит о том, что коннектор русского языка переведен именно коннектором французского языка, а не языковой единицей или конструкцией другой категории.
7. Рубрика «Dvrg» говорит о том, что коннектор русского языка переведен языковой единицей или конструкцией другой категории (не коннектором).
8. Рубрика «Повышение качества НМП» и другие рубрики, характеризующие изменение качества перевода, проставляются в последней, 12-й версии перевода, но характеризуют всю серию из 12 переводов.

Таблица 5 Серия аннотаций со снижением качества НМП

№ п/п	№ в НБД	Контекст РР в оригинале	РР в ори- гинале и ее при- знаки	Контекст РР в переводе	РР в переводе и ее признаки	При- знаки анно- тации	Дата и время записи в НБД
1	28213	Каримов сказал: — У нас уничтожили не только людей, национальную культуру уничтожили.	не только Ø (неединственности) (CNT p CNT q) (CNT) (Контакт)	Karimov a déclaré: « Pas seulement les gens ont été détruits ici, la culture nationale a été détruite».	pas seulement Ø (NoError)	Cngrn MT	01.03.19 15:18
2	28724			-// -	-// -	-// -	02.04.19 14:49
3	29413			-// -	-// -	-// -	06.05.19 14:45
4	30049			-// -	-// -	-// -	07.06.19 14:31
5	30340			-// -	-// -	-// -	05.07.19 14:28
6	30557			Karimov a déclaré: — Nous n'avons pas seulement détruit des personnes, nous avons également détruit la culture nationale.	ne pas seulement également (ErrorSemant)	-// -	05.08.19 22:29
7	30761			-// -	-// -	-// -	04.09.19 15:08
8	31181			-// -	-// -	-// -	06.10.19 23:28
9	31599			-// -	-// -	-// -	05.11.19 14:39
10	32186			-// -	-// -	-// -	03.12.19 15:08
11	32594			Karimov a déclaré: — Nous avons non seulement détruit des gens, nous avons détruit la culture nationale.	non seulement Ø (ErrorSemant)	-// -	04.01.20 11:51
12	33253			-// -	non seulement Ø (ErrorSemant) (Снижение ка- чества НМП)	-// -	05.02.20 22:40

В табл. 5 представлена серия переводов следующего русскоязычного фрагмента: «Каримов сказал: — У нас уничтожили не только людей, национальную культуру уничтожили» [В. С. Гроссман. Жизнь и судьба (ч. 1) (1959)]¹.

За год перевод исходного фрагмента также менялся 3 раза. Первая версия от 01.03.19 — приемлемый вариант перевода, поэтому для нее проставлена рубрика NoError. В 6-й версии от 05.08.19 преставлена рубрика ErrorSemant: в переводе искажен смысл исходного фрагмента из-за ошибки в выборе семантического субъекта (вместо «у нас уничтожили» получилось «мы уничтожили»). В 11-й версии перевода *des personnes* заменяется на *des gens*, но указанная ErrorSemant при этом сохраняется.

Для данной серии переводов преставлена рубрика, характеризующая снижение качества НМП, так как с течением времени в изначально допустимом варианте перевода появилась ошибка.

Отметим, что при выборе рубрик «Повышение/снижение качества НМП» оценивается общая динамика качества, при этом наличие в серии переводов допустимого варианта перевода (как в табл. 5) не обязательно: в табл. 4 перевод остается ошибочным, но его качество повышается за счет уменьшения числа ошибок.

5 Заключение

В статье приведено описание третьего этапа эксперимента по многократному фиксированию машинных переводов, полученных с помощью системы НМП Google. Непосредственной целью данного этапа был количественный анализ результатов, полученных в ходе исследования нестабильности НМП. В итоге было выведено и проанализировано соотношение категорий, характеризующих нестабильность НМП, и видов ошибок НМП. Также описана динамика ошибок на протяжении всего эксперимента и представлены данные распределения серий переводов по видам нестабильности.

Полученные данные о динамике ошибок свидетельствуют не только об актуальности изучения предельных возможностей обучения ИНС (что могут и чего не могут делать компьютерные системы на основе ИНС [17]), но и о необходимости индикаторной оценки результатов обучения при решении задач машинного перевода. Разработанная система индикаторов и их вычисленные значения уникальны. Предлагаемый подход к индикаторной оценке результатов обучения ИНС может позиционироваться как новый аспект в исследованиях качества машинного перевода (о необходимости таких исследований см., например, [18–20]).

¹Профессиональный перевод этого предложения на французский язык: «— Mais chez nous, reprit Karimov, on ne s'est pas contenté d'anéantir des hommes, on a anéanti toute une culture» [Alexis Berelowitch, 1980].

Литература

1. *Fenstermacher K. D.* The tyranny of tacit knowledge: What artificial intelligence tells us about knowledge representation // 38th Annual Hawaii Conference (International) on System Sciences Proceedings. — Washington, D.C., USA: IEEE Computer Society, 2005. Vol. 8. P. 243a. doi: 10.1109/HICSS.2005.620.
2. *Sako M.* Artificial intelligence and the future of professional work // Commun. ACM, 2020. Vol. 63. No. 4. P. 25–27.
3. *Denning P. J., Denning D. E.* Dilemmas of artificial intelligence // Commun. ACM, 2020. Vol. 63. No. 3. P. 22–24.
4. *Инькова-Манзотти О. Ю.* Коннекторы противопоставления во французском и русском языках. Сопоставительное исследование. — М.: Информэлектро, 2001. 429 с.
5. *Зализняк А. А., Зацман И. М., Инькова О. Ю., Кружков М. Г.* Надкорпусные базы данных как лингвистический ресурс // Корпусная лингвистика: Тр. VII Междунар. конф. — СПб.: СПбГУ, 2015. С. 211–218.
6. *Дурново А. А., Зацман И. М., Лощилова Е. Ю.* Кросслингвистическая база данных для аннотирования логико-семантических отношений в тексте // Системы и средства информатики, 2016. Т. 26. № 4. С. 124–137.
7. *Зализняк А. А., Зацман И. М., Инькова О. Ю.* Надкорпусная база данных коннекторов: построение системы терминов // Информатика и её применения, 2017. Т. 11. Вып. 1. С. 100–108.
8. *Зацман И. М., Кружков М. Г.* Надкорпусная база данных коннекторов: развитие системы терминов проектирования // Системы и средства информатики, 2018. Т. 28. № 4. С. 156–167.
9. *Бунтман Н. В., Гончаров А. А., Зацман И. М., Нуриев В. А.* Количественный анализ результатов машинного перевода с использованием надкорпусных баз данных // Информатика и её применения, 2018. Т. 12. Вып. 4. С. 96–105.
10. *Егорова А. Ю., Зацман И. М., Мамонова О. С.* Надкорпусные базы данных в лингвистических проектах // Системы и средства информатики, 2019. Т. 29. № 3. С. 77–91.
11. Национальный корпус русского языка. <http://www.ruscorpora.ru>.
12. *Егорова А. Ю., Зацман И. М., Косарик В. В., Нуриев В. А.* Нестабильность нейронного машинного перевода // Системы и средства информатики, 2020. Т. 30. № 2. С. 124–135.
13. *Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А.* Методика темпоральной оценки нестабильности машинного перевода // Системы и средства информатики, 2020. Т. 30. № 3. С. 67–80.
14. *Гончаров А. А., Бунтман Н. В., Нуриев В. А.* Ошибки в машинном переводе: проблемы классификации // Системы и средства информатики, 2019. Т. 29. № 3. С. 92–103.
15. *Инькова О. Ю.* Надкорпусная база данных как инструмент формальной вариативности коннекторов // Компьютерная лингвистика и интеллектуальные технологии: по мат-лам ежегодной Международ. конф. «Диалог». — М.: РГГУ, 2018. Вып. 17(24). С. 240–253.
16. *Добровольский Д. О., Кретов А. А., Шаров С. А.* Корпус параллельных текстов: архитектура и возможности использования // Национальный корпус русского языка: 2003–2005 / Отв. ред. В. А. Плунгян. — М.: Индрик, 2005. С. 263–296.

17. Brynjolfsson E., Mitchell T. What can machine learning do? Workforce implications // Science, 2017. Vol. 358. Iss. 6370. P. 1530–1534.
18. Bowker L., Buitrago Ciro J. Machine translation and global research: Towards improved machine translation literacy in the scholarly community. — Bingley, U.K.: Emerald Publishing, 2019. 111 p.
19. Burchardt A., Lommel A., Macketanz V. A new deal for translation quality // Universal Access Inf., 2020. doi: 10.1007/s10209-020-00736-5.
20. Koehn Ph. Neural machine translation. — New York, NY, USA: Cambridge University Press, 2020. 408 p.

Поступила в редакцию 14.09.20

MACHINE TRANSLATION: INDICATOR-BASED EVALUATION OF TRAINING PROGRESS IN NEURAL PROCESSING

A. Yu. Egorova, I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

Abstract: The paper presents data collected while observing training progress of a neural machine translation (NMT) engine. The observed training progress received qualitative evaluation based on a set of indicators. Two hundred and fifty text fragments in Russian were used as experimental material for the study. For the duration of one year, every month these fragments were translated into French using the publicly available Google’s NMT engine. The produced translations were recorded and annotated by language experts in a supracorpora database which resulted in a series of 12 annotated translations for each of the 250 Russian fragments. The annotations include labels of translation errors which enables researchers to determine the NMT instability types according to the changes of translation quality or lack thereof. The goal of this paper is to describe the newly developed indicator-based approach and to provide an example of its application to evaluation of a neural network training progress.

Keywords: neural machine translation; instability of machine translation; indicator-based evaluation; linguistic annotation; instability types

DOI: 10.14357/08696527200412

References

1. Fenstermacher, K. D. 2005. The tyranny of tacit knowledge: What artificial intelligence tells us about knowledge representation. *38th Annual Hawaii Conference (International) on System Sciences Proceedings*. Washington, D.C.: IEEE Computer Society. 8:243a. doi: 10.1109/HICSS.2005.620.

2. Sako, M. 2020. Artificial intelligence and the future of professional work. *Commun. ACM* 63(4):25–27.
3. Denning, P. J., and D. E. Denning. 2020. Dilemmas of artificial intelligence. *Commun. ACM* 63(3):22–24.
4. Inkova-Manzotti, O. Yu. 2001. *Konnektory protivopostavleniya vo frantsuzskom i russkom yazykakh. Sopostavitel'noe issledovanie* [Connectors of opposition in French and Russian: A comparative study]. Moscow: Informelektro. 429 p.
5. Zaliznyak, A. A., I. M. Zatsman, O. Yu. Inkova, and M. G. Kruzhkov. 2015. Nadkorpusnye bazy dannykh kak lingvisticheskiy resurs [Supracorpora databases as linguistic resource]. *Tr. VII Mezhdunar. Konf. "Korpusnaya lingvistika"* [7th Conference (International) "Corpus Linguistics" Proceedings]. St. Petersburg: SPbSU. 211–218.
6. Durnovo, A. A., I. M. Zatsman, and E. Yu. Loshchilova. 2016. Krosslingvisticheskaya baza dannykh dlya annotirovaniya logiko-semanticeskikh otnosheniy v tekste [Cross-lingual database for annotating logical-semantic relations in the text]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(4):124–137.
7. Zaliznyak, A. A., I. M. Zatsman, and O. Yu. Inkova. 2017. Nadkorpusnaya baza dannykh konnektorov: postroenie sistemy terminov [Supracorpora database on connectives: Term system development]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):100–108.
8. Zatsman, I. M., and M. G. Kruzhkov. 2018. Nadkorpusnaya baza dannykh konnektorov: razvitiye sistemy terminov [Supracorpora database of connectives: Design-oriented evolution of the term system]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(4):156–167.
9. Buntman, N. V., A. A. Goncharov, I. M. Zatsman, and V. A. Nuriev. 2018. Količestvennyy analiz rezul'tatov mashinnogo perevoda s ispol'zovaniem nadkorpusnykh baz dannykh [Using supracorpora databases for quantitative analysis of machine translations]. *Informatika i ee Primeneniya — Inform. Appl.* 12(4):96–105.
10. Egorova, A. Yu., I. M. Zatsman, and O. S. Mamonova. 2019. Nadkorpusnye bazy dannykh v lingvisticheskikh proektakh [Supracorpora databases in linguistic projects]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):77–91.
11. Natsional'nyy korpus russkogo yazyka [Russian National Corpus]. Available at: <http://www.ruscorpora.ru/> (accessed September 11, 2020).
12. Egorova, A. Yu., I. M. Zatsman, V. V. Kosarik, and V. A. Nuriev. 2020. Nestabil'nost' neyronnogo mashinnogo perevoda [Instability of neural machine translation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):124–135.
13. Egorova, A. Yu., I. M. Zatsman, M. G. Kruzhkov, and V. A. Nuriev. 2020. Metodika temporal'noy otsenki nestabil'nosti mashinnogo perevoda [The technique allowing for temporal estimation of the machine translation instability]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):67–80.
14. Goncharov, A. A., N. V. Buntman, and V. A. Nuriev. 2019. Oshibki v mashinnom perevode: problemy klassifikatsii [Machine translation errors: Problems of classification]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(3):92–103.
15. Inkova, O. Yu. 2018. Nadkorpusnaya baza dannykh kak instrument formal'noy variativnosti konnektorov [Supracorpora database as an instrument of the study of the formal variability of connectives]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii:*

- po mat-lam ezhegodnoy Mezhdunar. konf. “Dialog”* [Computer Linguistic and Intellectual Technologies: Conference (International) “Dialog” Proceedings]. Moscow. 17(24):240–253.
- 16. Dobrovolskiy, D. O., A. A. Kretov, and S. A. Sharov. 2005. Korpus parallel'nykh tekstov: arkhitektura i vozmozhnosti ispol'zovaniya [Corpus of parallel texts: Architecture and applications]. *Natsional'nyy korpus russkogo yazyka: 2003–2005* [Russian National Corpus: 2003–2005]. Moscow: Indrik. 263–296.
 - 17. Brynjolfsson, E., and T. Mitchell. 2017. What can machine learning do? Workforce implications. *Science* 358(6370):1530–1534.
 - 18. Bowker, L., and J. B. Ciro. 2019. *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Bingley, U.K.: Emerald Publishing. 111 p.
 - 19. Burchardt, A., A. Lommel, and V. Mackenz. 2020. A new deal for translation quality. *Universal Access Inf.* doi: 10.1007/s10209-020-00736-5.
 - 20. Koehn, Ph. 2020. *Neural machine translation*. New York, NY: Cambridge University Press. 408 p.

Received September 14, 2020

Contributors

Egorova Anna Yu. (b. 1991) — junior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ann.shurova@gmail.com

Zatsman Igor M. (b. 1952) — Doctor of Science in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; izatsman@yandex.ru

Kruzhkov Mikhail G. (b. 1975) — senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; magnit75@yandex.ru

Nuriev Vitaly A. (b. 1980) — Candidate of Science (PhD) in philology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; nurieff.v@gmail.com

МЕХАНИЗМ ТЕМПОРАЛЬНОГО СОПОСТАВЛЕНИЯ КОНКРЕТНО-ИСТОРИЧЕСКИХ ФАКТОВ

И. М. Адамович¹, О. И. Волков²

Аннотация: Статья продолжает серию работ, посвященных технологии поддержки конкретно-исторических исследований (ПКИИ). Технология построена на принципах сотворчества и краудсорсинга и ориентирована на широкий круг пользователей, не относящихся к профессиональным историкам и биографам. Статья посвящена дальнейшему развитию технологии путем интеграции в нее механизма темпорального сопоставления конкретно-исторических фактов. Приведены основные направления сопоставления фактов в конкретно-историческом исследовании, обоснована особая значимость подзадачи сопоставления и взаимоувязывания фактов по оси времени с целью определения последовательности событий в жизни объекта исследования и уточнения их датировки. Предложен подход к автоматизации этой подзадачи в рамках технологии, основанной на механизме автоматизированного поиска противоречий в конкретно-исторической информации. Показана возможность сведения задачи поиска допустимых временных интервалов к задаче линейного программирования. Обосновано применение симплекс-метода для реализации алгоритма темпорального сопоставления конкретно-исторических фактов.

Ключевые слова: конкретно-историческое исследование; распределенная технология; сопоставление фактов; историко-биографический факт; линейное программирование

DOI: 10.14357/08696527200413

1 Введение

Поддержка конкретно-исторических исследований стала одной из актуальных задач современности, что обусловлено вовлечением в исследовательский процесс не только членов профессионального исторического сообщества, но и самых широких слоев непрофессионалов в связи со все возрастающим интересом к частной, семейной истории [1].

В [2, 3] описана разработанная в ФИЦ ИУ РАН распределенная технология ПКИИ, основанная на принципах краудсорсинга (мобилизации ресурсов широкого круга добровольцев посредством информационных технологий). Данные в этой технологии организованы в форме семантической сети. Узлы сети представляют собой именованные универсальные классы объектов. Факты задаются

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Adam@amsd.com

²Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, Volkov@amsd.com

значениями экземпляров классов и связями между ними. Связи наследуются из сети классов [4].

Специфика конкретно-исторического исследования состоит в том, что факты, которыми оперирует исследователь (историко-биографические факты), обладают свойством фрагментарности и противоречивости. Получение этих фактов — необходимая и наиболее трудоемкая стадия биографического исследования, состоящая в работе в архивах и просмотре огромного числа документов. Работа исследователя на втором этапе состоит в сопоставлении выявленных фактов между собой с целью взаимоувязывания и разрешения противоречий [5].

2 Механизм автоматизированного поиска противоречий

В [6] было описано развитие технологии ПКИИ за счет включения в нее механизма, предназначенног для автоматизированного поиска противоречий в историко-биографических фактах.

Особенность поиска противоречий в историко-биографической информации при проведении конкретно-исторического исследования состоит в том, что биографические факты, увязывающие объекты исторической реальности с их динамическими характеристиками (атрибутами), связаны определенными зависимостями. Эти зависимости регламентируются естественными закономерностями, определяющими допустимые последовательности событий, задающими некий шаблон, сценарий или набор ограничений, законами природы, законами и нормативными документами, традициями. Формулировки регламентирующих закономерностей, законов, традиций называются нормальями. Именно нормали позволяют сопоставить факты, выявить и разрешить противоречия, сформировать новые факты-следствия, интегрировать данные в общую картину [7].

Эта особенность конкретно-исторического исследования в условиях технологии ПКИИ приводит к тому, что задача поиска прямого противоречия между двумя фактами становится тривиальной и сводится к сравнению разных значений одного атрибута одного объекта. Интерес для исследователя представляет поиск неформального противоречия между группой фактов и связывающими их нормальями. Этот интерес определяется тем, что благодаря специфике конкретно-исторического исследования выявление противоречия вместо сужения объема информации выявляет новые направления научного поиска и тем самым дает новый импульс исследованию.

Противоречия такого рода не могут быть разрешены с помощью традиционных подходов и требуют создания специального механизма. В [6] описан такой механизм: специфицированы соответствующие изменения объектной модели технологии за счет включения в нее понятия нормали и введения явного указания темпоральности атрибута и подробно описан алгоритм, в основе которого лежит опора на прецеденты — противоречия, явно выявленные пользователями. Выбор данного подхода объясняется тем, что структура и состав семантической сети технологии ПКИИ не определены заранее, а создаются динамически широким

кругом не связанных между собой исследователей, из чего следует, что алгоритм поиска противоречий не может опереться на заранее известное расположение нормалей в базе данных технологии, какие-либо изначально фиксированные их связи с атрибутами объектов, состав и наименование темпоральных атрибутов и т. п.

Алгоритм имеет три подсистемы, обеспечивающие последовательное выполнение следующих этапов:

- (1) создание прецедента пользователем;
- (2) автоматический поиск противоречий по прецеденту;
- (3) обработка пользователем результатов поиска.

3 Сопоставление фактов в конкретно-историческом исследовании

Как показано выше, задача исследователя в конкретно-историческом исследовании при обработке выявленной информации состоит помимо поиска противоречий в сопоставлении фактов с целью их взаимоувязывания. Специфика биографического исследования состоит в том, что в центре внимания исследователя находится конкретная личность и все без исключения стороны (социальные, экономические, политические, этнические, художественные и т. п.) ее реальной жизни [8]. Многообразие изучаемых аспектов жизни индивидуума, стоящего в центре исследования, приводит к появлению огромного числа подзадач сопоставления фактов. В качестве примеров таких подзадач, возникавших в реальных биографических исследованиях, можно привести:

- сопоставление сведений о передислокации полка с датами писем военнослужащего с целью определения города, в котором произошло упоминаемое в письме событие;
- сопоставление записей в метрических книгах и исповедных списках с целью выявления конкретных родов (очень актуальная и сложная задача, поскольку фамилии для крестьянства были введены только в XIX в., имена в разных документах могли писаться по-разному, а возраст мог сознательноискажаться, например с целью отсрочить призыв на воинскую службу в военное время);
- сопоставление Памятных книг губерний с информацией об изменении их границ и о переименовании деревень, сел, городов, волостей и уездов с целью определения расположения ныне не существующих населенных пунктов.

Среди многообразия аспектов, по которым проводится сопоставление, особняком стоит ось времени. Задача взаимоувязывания фактов о тех или иных событиях в жизни индивидуума с целью определения их последовательности и уточнения датировки является наиболее важной при проведении биографического исследования и, без преувеличения, лежит в его основе. Определению подхода к автоматизации именно этой задачи в рамках технологии ПКИИ и посвящена данная статья.

Для автоматизации задачи определения последовательности событий и уточнения их датировки можно опереться на описанный выше механизм автоматического поиска противоречий в конкретно-исторической информации. Основной принцип состоит в следующем: допустимая привязка группы фактов к оси времени не должна вызывать противоречий между ними и связывающими их нормальными, что, в свою очередь, может быть определено автоматически. Ниже приведен пример задачи, возникшей при проведении биографического исследования, которая может быть решена в рамках данного подхода.

1. Известные факты из жизни изучаемой персоны:

- дата рождения — после 1857 г.;
- по данным исповедной росписи 1892 г. проживал с женой и сыном 5 лет от роду, числясь при этом отставным старшим унтер-офицером 15-го пехотного Шлиссельбургского полка.

2. Соответствующие нормали:

- Устав о воинской повинности от 1 января 1874 г. гласит: «Поступление на службу по призывам решается жребием... К жребию призываются ежегодно один только возраст населения, именно молодые люди, которым к 1-му января того года, когда набор производится, минуло двадцать лет от роду. Общий срок службы в сухопутных войсках для поступающих по жребию определяется в пятнадцать лет, из коих шесть лет действительной службы и девять лет в запасе»;
- социальная норма того времени определяет естественным рождение ребенка после заключения брака;
- брак для нижних чинов в принципе допускался с разрешения начальства, но фактически был возможен только после увольнения в запас;
- биологическая норма определяет нормальный срок беременности в 9 месяцев.

Сопоставление временных интервалов позволяет выявить следующее ограничение: объект изучения родился, скорее всего, не позже 1860 г. При этом исследователю следует иметь в виду, что возможны обстоятельства, при которых данные ограничения могут не действовать: сокращенный срок службы, рождение ребенка вне брака и т. п. Эти обстоятельства должны быть проверены в рамках исследования.

4 Алгоритм сопоставления фактов

Работа алгоритма начинается с указания пользователем объекта (элемента сети экземпляров), для которого требуется уточнение датировок связанных с ним событий. Алгоритм ищет темпоральные атрибуты $T = \{\tau_i\}$ с неустановленными

значениями для этого объекта и для объектов, с атрибутами которых заданный пользователем объект связан связью «значение», т. е. выявляет события без датировки, оформленные как атрибуты данного объекта и как самостоятельные сущности, к которым этот объект привязан как участник [2].

Задача сопоставления состоит в нахождении таких возможных граничных значений выявленных темпоральных атрибутов T , при которых не возникает противоречий, механизм проверки которых определен [6] и базируется на precedентах, заданных пользователями. При этом для каждого precedента фиксируются:

- набор условий соответствия нормалям;
- набор условий противоречия.

Условие противоречия представляет собой ссылку на нормаль и логико-арифметическое выражение над значениями атрибутов неких объектов, истинность которого является требованием нормали, но не выполняется в реальности для данных объектов, в чем и проявляется противоречие.

Условие соответствия нормали определяет применимость нормали к данной ситуации. Условия разделяются на темпоральные и прочие. Обязательное темпоральное условие определяет соответствие даты основного события срокам действия нормали. Такое условие задается ссылкой на нормаль и значением темпорального атрибута. Прочие условия задаются аналогично условиям противоречия, т. е. ссылкой на нормаль и логико-арифметическим выражением над значениями атрибутов неких объектов, истинность которого служит критерием применимости нормали.

Для каждого набора значений выявленных темпоральных атрибутов, для которых исходно значения не заданы, механизмом поиска противоречий может быть проверено наличие или отсутствие противоречий с помощью соответствующего механизма [6]. При наличии противоречий набор значений будем называть противоречивым, в противном случае — непротиворечивым. Цель работы описываемого алгоритма — выявление минимального (с минимальными значениями) и максимального (с максимальными значениями) непротиворечивых наборов, т. е. определение границ допустимых значений темпоральных атрибутов T .

Строить алгоритм на переборе всех возможных значений выявленных темпоральных атрибутов с целью определения допустимых границ невозможно, поскольку проверяемые диапазоны должны быть по длительности сравнимы со сроком человеческой жизни (десятки лет), и при наличии около десятка лакун в датировках событий из жизни объекта исследования (что по опыту реальных исследований служит вполне правдоподобной оценкой), объем перебора составит более 10^{35} вариантов. Поэтому при построении алгоритма целесообразно опереться на специфическую структуру условий в описании precedентов из наборов условий соответствия нормалям и наборов условий противоречия.

Каждое условие является логико-арифметическим выражением над значениями атрибутов объектов, т. е. логическим выражением, которое без ограничения

общности можно считать представленным в дизъюнктивной нормальной форме (ДНФ), т. е. в форме дизъюнкции конъюнкций литералов, в качестве литералов которого выступают арифметические условия на значения атрибутов. Эти арифметические условия представляют собой неравенство или равенство с константой в правой части и арифметическим выражением над значениями атрибутов в левой. По опыту реальных конкретно-исторических исследований и исходя из здравого смысла можно считать, что условия на темпоральные атрибуты объектов исследования могут задавать либо требование вхождения даты в определенный диапазон, либо требование вхождения разницы двух дат в определенный диапазон. Поэтому для темпоральных атрибутов левая часть арифметического условия может быть либо значением атрибута, либо разницей значений двух атрибутов.

Арифметические условия, не зависящие от значений выявленных темпоральных атрибутов T , можно считать логическими константами, поскольку они вычисляются однократно. Следовательно, условия соответствия нормалям и условия противоречия по определению ДНФ и в соответствии с описанным выше видом арифметических условий, содержащих темпоральные атрибуты $\tau_i \in T$, могут быть представлены в форме совокупности систем линейных неравенств или уравнений, и задача сводится к нахождению $\max \tau_i$ и $\min \tau_i$ при ограничениях, заданных системами линейных относительно T уравнений или неравенств. При такой постановке задача сводится к задаче линейного программирования в общей форме: нахождение наибольшего (наименьшего) значения линейной функции многих переменных при линейных ограничениях типа равенств (или неравенств), когда на переменные есть (или нет) ограничения на знак [9].

Методы решения задач линейного программирования разработаны и подробно изложены в соответствующей литературе [10]. На практике высокую эффективность показал симплекс-метод [11], который и предлагается использовать для решения задачи сопоставления конкретно-исторических фактов.

5 Выводы

Предложенный механизм автоматизированного сопоставления историко-биографических фактов существенно дополняет и развивает технологию ПКИИ, ориентированной на широкий круг не являющихся профессиональными историками и биографами пользователей, что очень актуально в связи со все возрастающим общественным интересом к частной, семейной истории.

Данный механизм помимо решения своей основной задачи определения последовательности событий и уточнения их датировки расширяет возможности подсистемы поиска противоречий, поскольку позволяет выявить темпоральные противоречия в группе фактов, даже если отсутствует их точная привязка к оси времени. В этом случае подсистема поиска противоречий неприменима, но построенный на ее базе механизм сопоставления фактов может быть способен определить, что допустимые временные интервалы для рассматриваемых фактов пусты, из чего будет следовать, что интерпретация исследователем событий из

жизни объекта исследования, их последовательности и взаимосвязи неверна, и ему следует сконцентрировать свои усилия на поиске дополнительной информации, способной дать исследованию новый импульс. Такие ситуации часто возникают в реальных конкретно-исторических исследованиях, и инструмент для их автоматизированной обработки очень востребован.

Литература

1. Помникова А. Ю. Семейная история в дискурсивном пространстве // Вестник Мининского университета, 2019. Т. 7. № 1. С. 9.
2. Адамович И. М., Волков О. И. Технология распределенного автоматизированного анализа исторических текстов // Системы и средства информатики, 2016. Т. 26. № 3. С. 148–161.
3. Адамович И. М., Волков О. И. Единая технология поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 194–205.
4. Адамович И. М., Волков О. И. Принципы организации данных для технологии поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 2. С. 161–171.
5. Адамович И. М., Волков О. И. Иерархическая форма представления биографического факта // Системы и средства информатики, 2016. Т. 26. № 2. С. 108–122. doi: 10.14357/08696527160207.
6. Адамович И. М., Волков О. И. Автоматизированный поиск противоречий в конкретно-исторической информации // Системы и средства информатики, 2020. Т. 30. № 3. С. 145–153. doi: 10.14357/08696527200313.
7. Маркова Н. А. Логика биографических фактов // Информатика и её применения, 2012. Т. 6. Вып. 2. С. 87–96.
8. Иконникова С. Н. Биографика как часть исторической культурологии // Вестник СПбГУКИ, 2012. № 2(11). С. 6–10.
9. Зенкевич Н. А., Губар Е. А. Практикум по исследованию операций. — СПб.: Золотое сечение, 2007. 170 с.
10. Карманов В. Г. Математическое программирование. — 5-е изд. — М.: Физматлит, 2004. 264 с.
11. Панюков А. В., Горбик В. В. Параллельные реализации симплекс-метода для безошибочного решения задач линейного программирования // Вестник ЮУрГУ. Серия: Математическое моделирование и программирование, 2011. № 25(242). С. 107–118.

Поступила в редакцию 08.05.20

MECHANISM OF TEMPORAL COMPARISON OF CONCRETE-HISTORICAL FACTS

I. M. Adamovich and O. I. Volkov

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The article continues the series of works devoted to the technology of concrete historical research supporting. The technology is based on the principles of co-creation and crowdsourcing and is designed for a wide range of users which are not professional historians and biographers. The article is devoted to the further development of the technology by integrating the mechanism of temporal comparison of concrete-historical facts. The main directions of facts comparison in concrete-historical research are given. The high significance of the subtask of facts comparison and linking along the time axis in order to determine the sequence of events in the life of the object of research and clarify their dating is established. The approach to automation of this subtask in the context of the technology based on the mechanism for the automated search for contradictions in concrete-historical information is proposed. The possibility of reducing the problem of finding admissible time intervals to the linear programming problem is shown. The use of the Simplex algorithm for implementation of the algorithm of temporal comparison of concrete-historical facts is reasoned.

Keywords: concrete historical investigation; distributed technology; facts comparison; historical-biographical fact; linear programming

DOI: 10.14357/08696527200413

References

1. Pomnikova, A. Yu. 2019. Semeynaya istoriya v diskursivnom prostranstve [Family stories in different types of discourse]. *Vestnik Mininskogo universiteta* [Minin University Bull.] 7(1):9.
2. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelennogo avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):148–161. doi: 10.14357/08696527160311.
3. Adamovich, I. M., and O. I. Volkov. 2019. Edinaya tekhnologiya podderzhki konkretno-istoricheskikh issledovanii [Unified technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):194–205. doi: 10.14357/08696527190116.
4. Adamovich, I. M., and O. I. Volkov. 2019. Printsipy organizatsii dannykh dlya tekhnologii podderzhki konkretno-istoricheskikh issledovanii [The principles of data organization for the technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 2(29):161–171. doi: 10.14357/08696527190214.

5. Adamovich, I. M., and O. I. Volkov. 2016. Ierarkhicheskaya forma predstavleniya biograficheskogo fakta [Hierarchical format of biographical fact]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 2(26):108–122. doi: 10.14357/08696527160207.
6. Adamovich, I. M., and O. I. Volkov. 2020. Avtomatizirovannyy poisk protivorechiy v konkretno-istoricheskoy informatsii [Automated search for contradictions in concrete-historical information]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):145–153. doi: 10.14357/08696527200313.
7. Markova, N. A. 2012. Logika biograficheskikh faktov [A logic of biographical facts]. *Informatika i ee Primeneniya — Inform. Appl.* 6(2):87–96.
8. Ikonomikova, S. N. 2012. Biografika kak chast' istoricheskoy kul'turologii [Biographical studies as part of the historical cultural studies]. *Vestnik SPbGU* [Bull. of Saint-Petersburg State University of Culture and Art] 2(11):6–10.
9. Zenkevich, N. A., and E. A. Gubar. 2007. *Praktikum po issledovaniyu operatsiy* [Lessons on operations research]. St. Petersburg: Zolotoe sechenie. 170 p.
10. Karmanov, V. G. 2004. *Matematicheskoe programmirovaniye* [Mathematical programming]. 5th ed. Moscow: Fizmatlit. 264 p.
11. Panyukov, A. V., and V. V. Gorbik. 2011. Parallel'nye realizatsii simpleks-metoda dlya bezshibochnogo resheniya zadach lineynogo programmirovaniya [The parallel simplex-method achievements for errorless solving of linear programming problems]. *Vestnik YUUrGU. Ser. Matematicheskoe modelirovanie i programmirovaniye* [Bull. of the South Ural State University. Mathematical modelling, programming & computer software ser.] 242(25):107–118.

Received May 8, 2020

Contributors

Adamovich Igor M. (b. 1934) — Candidate of Science (PhD) in technology, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Adam@amsd.com

Volkov Oleg I. (b. 1964) — leading programmer, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Volkov@amsd.com

БЛОКЧЕЙН-ТЕХНОЛОГИИ НОРМАЛИЗОВАННОГО БЮДЖЕТНОГО ОБЕСПЕЧЕНИЯ НАЦИОНАЛЬНЫХ ПРОЕКТОВ

А. В. Ильин¹, В. Д. Ильин²

Аннотация: Блокчейн-технологии нормализованного бюджетного обеспечения (НБО-технологии) рассматриваются как инструментальные средства, необходимые для ресурсного обеспечения разработки и исполнения национальных проектов защиты и развития потенциала страны. Рассматриваемая в статье часть методологического обеспечения разработок НБО-технологий включает методологии ситуационного онлайн-бюджетирования и нормализованного товарно-денежного обращения. Онлайн-планирование бюджета рассматривается как задача интервального планирования расходов с учетом ситуационно зависимых обязательных и ориентирующих требований к исковому решению. Задача решается методом целевого перемещения решения в режиме вычислительного эксперимента. Методология нормализованного товарно-денежного обращения включает методы формирования и реализации договорных отношений в цифровой среде, платежного и товарного кредитования, онлайн-банкинга на основе банков-провайдеров, корпоративных и персональных электронных банков. Блокчейн-технологии нормализованного бюджетного обеспечения практически исключают незапланированное использование бюджетных средств.

Ключевые слова: блокчейн-технологии нормализованного бюджетного обеспечения (НБО-технологии); ресурсное обеспечение национальных проектов защиты и развития потенциала страны; ситуационное онлайн-бюджетирование; нормализованное товарно-денежное обращение

DOI: 10.14357/08696527200414

1 Введение

Национальные проекты (нацпроекты) [1] разрабатываются и реализуются, чтобы обеспечить защиту и развитие *потенциала страны* [2]. В России задачами разработки, ресурсного обеспечения, исполнения и контроля результативности нацпроектов занимаются федеральные министерства и ведомства (Минэконом-

¹Государственный научно-исследовательский институт авиационных систем, ilyin@res-plan.com

²Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, vdilyin@yandex.ru

развития¹, Минфин², Казначейство³ и Счетная палата⁴) во взаимодействии с органами власти субъектов Российской Федерации.

Одной из задач ресурсного обеспечения нацпроектов является *задача бюджетного обеспечения*.

Блокчейн-технологии в арсенале современных технологий цифровой экономики [3–10]. В России блокчейн-технологии наиболее успешно применяются на предприятиях энергетики, добывающей и обрабатывающей промышленности, в финансовых учреждениях и логистике. Основные области применения: контроль поставок продукции; обмен документацией при сделках. Интенсивно увеличивается число блокчейн-проектов с целью: рационализации корпоративного документооборота; *токенизации*⁵ продукции и финансовых инструментов; безопасных платежей; контроля поставок. Создается блокчейн-инфраструктура рынка интеллектуальной собственности [11].

Запись формул и выделение фрагментов текста. Для выделения определений, замечаний и примеров используются средства языка TSM-комплекса (TSM: textual symbolic modeling), разработанного для формализованного описания текстовых моделей [12].

В статье применены следующие средства выделения фрагментов текста:
□ ⟨фрагмент описания⟩ □ ≈ утверждение (определение, аксиома и др.) (здесь и далее символ ≈ заменяет слово «означает»);

◊ ⟨фрагмент описания⟩ ◊ ≈ замечание;
○ ⟨фрагмент описания⟩ ○ ≈ пример.

Курсивом выделены первые вхождения названий понятий и фрагменты описания, к которым авторы хотят привлечь внимание.

Обсуждаемые результаты. Представлена часть результатов методологического обеспечения разработки блокчейн-технологий государственного планирования и управления процессами защиты и развития потенциала страны. Результаты получены при выполнении научно-исследовательской работы «Моделирование социальных, экономических и экологических процессов» (№ 0063-2016-0005), выполняемой в соответствии с государственным заданием ФАНО России для Федерального исследовательского центра «Информатика и управление» РАН. Они служат тематическим дополнением и развитием результатов исследований, посвященных *ситуационной информатизации жизнедеятельности населения страны*, рассматриваемой как совокупность образовательной, производственной и других видов деятельности [13].

¹ Министерство экономического развития Российской Федерации. Приоритетные направления. <https://www.economy.gov.ru/material/directions>.

² Министерство финансов Российской Федерации. Публичные цели. <https://minfin.gov.ru/ru/?fullversion=1>.

³ Казначейство России. <http://www.roskazna.ru/o-kaznachejstve>.

⁴ Счетная палата Российской Федерации. <https://ach.gov.ru/about>.

⁵ Токен — уникальный шифрованный код, соответствующий некоторому объекту.

◊ Материал статьи адресован прежде всего исследователям, занимающимся методологическим обеспечением разработок *информационных технологий цифровой экономики*. ◊

2 Основы концепции нормализованного бюджетного обеспечения нацпроектов

□ *Потенциал страны* как ресурсное основание ее обороноспособности и социально-экономического развития определяется характеристиками *населения, среды обитания и систем организации и обеспечения жизни*. □

Изменение потенциала страны позволяет судить о состоянии объекта государственного управления (одной из форм которого является реализация нацпроектов). ◊ Приоритетный список показателей уровня развития потенциала страны и их истолкование (включающее способы оценки и стимулирования изменений) составляют основу государственного планирования и управления процессами развития потенциала страны [2]. ◊

Блокчейн-технологии нормализованного бюджетного обеспечения, позволяющие повысить эффективность использования бюджетных средств и исключить их незапланированное расходование, представляют собой *блокчейн-реализацию методологии нормализованного товарно-денежного обращения*. ◊ Эта методология включает *методы формирования и реализации договорных отношений в цифровой среде* [14], *платежного и товарного кредитования* [15], *онлайн-банкинга на основе банков-провайдеров, корпоративных и персональных электронных банков* [5, 9, 10]. ◊

Совершенствование технологий ресурсного обоснования нацпроектов. Повышение эффективности нацпроектов предполагает совершенствование технологий их ресурсного обоснования. Среди них — технология *планирования расходов на разработку и реализацию проекта* (далее — *планирование расходов на проект*).

План расходов обычно обновляется в течение всего жизненного цикла проекта. Это связано с конкретизацией деталей проекта, уточнением величины располагаемых денежных средств, изменением цен и другими факторами. □ *Невозможность точного прогноза расходов по составляющим проекта — аксиома, справедливая и для его доходной части.* □

Более адекватным выглядит представление данных и результата в виде числовых отрезков (левая граница отрезка соответствует наименьшему ожидаемому значению, а правая — наибольшему) с *учетом реальной точности данных и реализации решений* [16].

Указанный методологический недостаток устранен в рассматриваемой далее обновленной постановке задачи *ситуационного планирования расходов*.

Оценка состояния доходной и расходной части проекта. □ *Ситуацией бюджетного обеспечения нацпроекта* (БОН-ситуацией) назовем состояние

доходной и расходной части проекта, представленное совокупностью параметров, описывающих источники поступления денежных средств, расходные статьи и др.

Целевой назовем БОН-ситуацию, которую планируется создать в результате завершения проекта; стартовой — существующую в начале разработки проекта; достигнутой — созданную в результате выполнения этапа проекта. □

□ Портрет БОН-ситуации — формализованное описание состояния доходной и расходной части проекта. □

Портрет ситуации формируется в соответствии со специальной системой правил [17]. На вход этой системы поступают данные о параметрах пространства состояний бюджетного обеспечения нацпроекта, о располагаемых типах воздействий и о ресурсах, обеспечивающих реализацию воздействий.

Динамика изменения ситуаций отражается посредством временных рядов, состоящих из портретов ситуаций. Это позволяет проследить последствия реализации воздействий. Портреты ситуаций хранятся в специальном архиве системы ресурсного обеспечения нацпроекта.

В рассматриваемой технологии ситуационного планирования расходов на нацпроекты мониторинг выполняется комплексом цифровых двойников [18], представляющим собой блокчейн-сетевую обучаемую символическую модель доходной и расходной части проекта, реализованную в человеко-машинной среде решения задач (*s-среде*) [12].

Ситуационная модель универсального договора в цифровой среде [14] используется для формирования и исполнения договоров, реализуемых при бюджетном обеспечении нацпроектов.

Блокчейн-технологии бюджетного обеспечения нацпроектов обеспечивают децентрализацию, прозрачность и безопасность процессов планирования, исполнения и контроля: повышается безопасность платежных операций, контролируемость расходов, исключаются потери электронных документов и др.

2.1 Ситуационное планирование расходов на нацпроект

Задача ситуационного планирования расходов на проект решается в режиме вычислительного эксперимента сообществом экспертов по ресурсному обеспечению нацпроектов. Их компьютеры (и/или компьютерные устройства) принадлежат множеству узлов блокчейн-сети *Обеспечивающей стороны*, готовящей проекты планов, поступающих на утверждение *Решающей стороной* [2].

Для числового отрезка $[a^{\min}, a^{\max}]$ ($a^{\min} \geq 0, a^{\max} > 0$), задающего распределенную величину (\bigcirc сумму денег \bigcirc), отрезков $[b_i^{\min}, b_i^{\max}]$ ($b_i^{\min} \geq 0, b_i^{\max} > 0, i = 1, \dots, n$), задающих запросы по расходным статьям, и весовых коэффициентов (приоритетов) расходных статей $p_i > 0$ ($i = 1, \dots, n$) требуется найти план расходов по статьям:

$$[x_i^{\min}, x_i^{\max}] : \left\{ 0 \leq x_i^{\min} \leq b_i^{\min}, x_i^{\max} \leq b_i^{\max}, \sum x_i^{\min} \leq a^{\min}, \sum x_i^{\max} \leq a^{\max}. i = 1, \dots, n \right\}. \quad (1)$$

Для совокупного вектора искомого плана $\mathbf{x} = (x_1^{\min}, \dots, x_n^{\min}, x_1^{\max}, \dots, x_n^{\max})$ может быть также задан набор требований

$$\mathbf{C}\mathbf{x} \leq \mathbf{d} \leftarrow \mathbf{q}, \quad (2)$$

где \mathbf{C} — матрица вещественных коэффициентов размера $k \times 2n$ ($k \geq 1$); \mathbf{d} — вектор-столбец вещественных констант размера k ; \mathbf{q} — вектор-столбец весовых коэффициентов (приоритетов) требований ($0 < q_i \leq +\infty, i = 1, \dots, k$).

Обязательные требования имеют приоритет $+\infty$. Приоритеты *ориентирующих требований* задаются положительными вещественными числами. \diamond Приоритеты задают эксперты-экономисты, учитывая относительную важность выполнения рассматриваемых требований. \diamond

Требования (1) являются обязательными. Требования (2) могут быть как обязательными, так и ориентирующими.

В трех следующих случаях задача решается итеративным методом *приоритетного интервального распределения* [19, 20], реализованным в действующем интернет-сервисе планирования расходов¹.

1. $\sum b_i^{\min} > a^{\min}, \sum b_i^{\max} > a^{\max}$ ($i = 1, \dots, n$) и набор требований (2) имеет вид:

$$\begin{aligned} \left\{ \sum x_i^{\min} = a^{\min} \leftarrow +\infty, \sum x_i^{\max} = a^{\max} \leftarrow +\infty \right. & (i = 1, \dots, n), \\ p_j b_j^{\min} x_i^{\min} - p_i b_i^{\min} x_j^{\min} = 0 \leftarrow 1 & \end{aligned}$$

для каждой пары i, j $\{1 \leq i \leq n, 1 \leq j \leq n\}$,

$$p_j (b_j^{\max} - b_j^{\min}) (x_i^{\max} - x_i^{\min}) - p_i (b_i^{\max} - b_i^{\min}) (x_j^{\max} - x_j^{\min}) = 0 \leftarrow 1$$

для каждой пары i, j : $\{b_i^{\max} > b_i^{\min}, b_j^{\max} > b_j^{\min}, 1 \leq i \leq n, 1 \leq j \leq n\}$,

$$p_j b_j^{\max} x_i^{\max} - p_i b_i^{\max} x_j^{\max} = 0 \leftarrow 1$$

- для каждой пары i, j : $\{b_i^{\max} = b_i^{\min}, b_j^{\max} = b_j^{\min}, 1 \leq i \leq n, 1 \leq j \leq n\}$. (3)

2. $\sum b_i^{\min} \leq a^{\min}, \sum b_i^{\max} > a^{\max}$ ($i = 1, \dots, n$) и набор требований (2) имеет вид:

¹Ильин А. В. Интернет-сервис планирования расходов. <https://www.res-plan.com/services-ru>.

$$\begin{aligned} & \left\{ x_i^{\min} = b_i^{\min} \leftarrow +\infty, \sum x_i^{\max} = a^{\max} \leftarrow +\infty \ (i = 1, \dots, n), \right. \\ & p_j (b_j^{\max} - b_j^{\min}) (x_i^{\max} - x_i^{\min}) - p_i (b_i^{\max} - b_i^{\min}) (x_j^{\max} - x_j^{\min}) = 0 \leftarrow 1 \\ & \text{для каждой пары } i, j: \left\{ b_i^{\max} > b_i^{\min}, b_j^{\max} > b_j^{\min}, 1 \leq i \leq n, 1 \leq j \leq n \right\}, \\ & p_j b_j^{\max} x_i^{\max} - p_i b_i^{\max} x_j^{\max} = 0 \leftarrow 1 \\ & \left. \text{для каждой пары } i, j: \left\{ b_i^{\max} = b_i^{\min}, b_j^{\max} = b_j^{\min}, 1 \leq i \leq n, 1 \leq j \leq n \right\} \right\}. \quad (4) \end{aligned}$$

3. $\sum b_i^{\min} > a^{\min}$, $\sum b_i^{\max} \leq a^{\max}$ ($i = 1, \dots, n$) и набор требований (2) имеет вид:

$$\begin{aligned} & \left\{ \sum x_i^{\min} = a^{\min} \leftarrow +\infty, x_i^{\max} = b_i^{\max} \leftarrow +\infty \ (i = 1, \dots, n), \right. \\ & p_j b_j^{\min} x_i^{\min} - p_i b_i^{\min} x_j^{\min} = 0 \leftarrow 1 \\ & \left. \text{для каждой пары } i, j \ \{1 \leq i \leq n, 1 \leq j \leq n\} \right\}. \quad (5) \end{aligned}$$

◊ *Прикладная точность*, задаваемая экспертом-экономистом, определяет в выбранных единицах измерения минимальную значимую сумму (○ целые степени 10 от -4 до 9 ○). Данные округляются в соответствии с заданной прикладной точностью. Результаты расчетов округляются по специальному алгоритму — с сохранением требуемой суммы, учетом запросов и приоритетов. Возможность задания прикладной точности позволяет, в частности, решать целочисленные задачи. ◊

В случаях когда набор требований (2) отличается от вариантов (3)–(5), задача решается методом целевого перемещения решения [21]. Для удобства читателя далее приведено схематическое описание идеи этого метода.

При несовместности системы требований в качестве начального плана рассматривается чебышевская точка [21], минимизирующая максимальный дефицит:

$$\begin{aligned} & \min_x \max_i (c_{i1}x_1 + \dots + c_{i,2n}x_{2n} - d_i), \quad i = 1, \dots, k \\ & \left(\text{здесь и далее для удобства записи } x_1 = x_1^{\min}, \dots, x_n = x_n^{\min}, x_{n+1} = x_1^{\max}, \dots, \right. \\ & \quad \left. \dots, x_{2n} = x_n^{\max} \right). \end{aligned}$$

Если система требований совместна, то начальным планом может служить как чебышевская точка, так и решение классической задачи линейного программирования.

На последующих шагах диалогового поиска плана эксперт-экономист, сопоставляя значения левых и правых частей набора неравенств $Cx \leq d$, оценивает

реализуемость и эффективность полученных решений. Текущий план, получивший оценку эксперта, может быть определен как окончательный или могут быть изменены требования, направляющие перемещение решения. Каждый план (с оценкой) может быть сохранен в базе данных для сравнительного анализа и возможного возврата.

Если \mathbf{x}^0 — текущий план и эксперт определил ориентирующие требования для перемещения от \mathbf{x}^0 к целевому плану \mathbf{x} в виде набора

$$\mathbf{C}^* \mathbf{x} = \mathbf{C}^* \mathbf{x}^0 + \mathbf{h},$$

где \mathbf{C}^* — подмножество из m строк матрицы \mathbf{C} , а $\mathbf{h} = (h_1, \dots, h_m)$ — вектор ненулевых констант, направляющих перемещение решения, то кандидатом на улучшенное решение будет:

$$\mathbf{x} = \left(x_1^0 + \frac{\sum q_i \Delta_i x_1}{\sum q_i} \quad (i = 1, \dots, m), \quad x_{2n}^0 + \frac{\sum q_i \Delta_i x_{2n}}{\sum q_i} \quad (i = 1, \dots, m) \right),$$

где

$$\Delta_i x_j = \frac{c_{ij} h_i}{c_{i1}^2 + \dots + c_{i,2n}^2}, \quad j = 1, \dots, 2n, \quad i = 1, \dots, m.$$

Такое решение «ближе» к гиперплоскостям, определяемым требованиями с более высокими приоритетами. Детальное обоснование метода и изложение алгоритма приведено в [21].

2.2 Применение технологий платежного и товарного кредитования

Технологии платежного и товарного кредитования рассчитаны прежде всего на увеличение информационной неуязвимости *товарно-денежного обращения в цифровой среде* [15].

Платежное кредитование нацпроекта рассматривается как программно контролируемая частичная или полная предоплата заказанных товаров, правила реализации которой жестко связаны с правилами исполнения заказа. График перечисления платежного кредита реализуется по *технологии назначенных платежей* [22]. *Товарное кредитование нацпроекта* — программно контролируемая поставка заказанных товаров с частичной или полной отсрочкой оплаты, правила реализации которой жестко связаны с правилами поставки (основано на *технологии электронной долговой торговли* [23]). Обеспечивается оформление отсроченной части оплаты товара как долга заказчика поставщику. Неоплаченная часть стоимости товара оформляется как долг, график выплаты которого фиксируется в договоре (где указываются штрафные санкции за нарушение графика выплат долга и ненадлежащее качество поставленного товара). Мониторинг платежного и товарного кредитования нацпроекта — задачи, решаемые

соответствующими комплексами цифровых двойников. Правила исполнения каждого договора контролируются программно во взаимодействии с цифровыми двойниками, отслеживающими события, определенные в договоре.

Поскольку технологии платежного и товарного кредита позволяют кредиторам, имеющим *персональные электронные банки* и/или *корпоративные электронные банки* [5, 9, 10], использовать только собственные средства, реализация этих технологий не связана с уменьшением товарной емкости денег, а долговая денежная эмиссия служит средством безинфляционного регулирования общей суммы денег в экономической системе.

3 Заключение

Бюджетное обеспечение нацпроектов — комплексная задача, от эффективности решения которой зависит потенциал страны, рассматриваемый как ресурсное основание ее обороноспособности и социально-экономического развития.

Блокчейн-технологии нормализованного бюджетного обеспечения гарантируют децентрализацию, прозрачность и безопасность процессов планирования и исполнения бюджетного обеспечения нацпроектов.

Технология ситуационного планирования расходов позволяет на всех этапах жизненного цикла нацпроекта получать решения, эффективность которых обеспечена соответствием *обязательным и ориентирующими требованиям*.

Реализация НБО-технологий предполагает необходимость организационно-технологического совершенствования бюджетных учреждений, занимающихся государственными проектами и ресурсным обеспечением их исполнения.

Литература

1. Национальные проекты: целевые показатели и основные результаты. — М., 2019. <http://static.government.ru/media/files/p7nn2CS0pVhvQ98OwAt2dzClAietQih.pdf>.
2. Ильин В. Д. Основания ситуационной информатизации. — М.: Наука, Физматлит, 1996. 180 с.
3. Tapscott D. The digital economy: Promise and peril in the age of networked intelligence. — New York, NY, USA: McGraw-Hill, 1996. 342 p.
4. Christensen C. M. The innovator's dilemma: When new technologies cause great firms to fail. — Boston, MA, USA: Harvard Business School Press, 1997. 288 p.
5. Ilyin A. V., Ilyin V. D. Towards a normalized economic mechanism based on E-services // Agris on-line Papers in Economics and Informatics, 2014. Vol. 6. Iss. 3. P. 39–49.
6. The New Digital Economy: How it will transform business. — Oxford Economics, 2015. 34 p.
7. G20 digital economy development and cooperation initiative // G20 Summit, 2016. <http://en.kremlin.ru/supplement/5111>.

8. Цифровая экономика Российской Федерации: Программа, утвержденная распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р.
9. *Ilyin A. V., Ilyin V. D.* The normalized economic mechanism in the digital environment // Int. J. Open Information Technologies, 2019. Vol. 7. Iss. 12. P. 77–83.
10. Ильин А. В., Ильин В. Д. Нормализованный экономический механизм: цифровые технологии поливалютного рынка // Системы и средства информатики, 2020. Т. 30. № 1. С. 186–197. doi: 10.14357 / 08696527200116.
11. Компания MINDSMITH провела масштабное исследование отечественного блокчейн-рынка // NBJ, 2019. <http://nbj.ru/pubs/banki-i-biznes/2019/11/26/kompanija-mindsmith-provela-masshtabnoe-issledovanie-otechestvennogo-blokchein-rynska/?full>.
12. Ильин В. Д. Символьное моделирование // Большая российская энциклопедия, 2018. http://dev.bigenc.ru/technology_and_technique/text/4010980.
13. Ильин А. В., Ильин В. Д. Основы концепции ситуационной информатизации жизнедеятельности // Системы и средства информатики, 2017. Т. 27. № 3. С. 188–201. doi: 10.14357 / 08696527170315.
14. Ильин А. В., Ильин В. Д. Ситуационная модель универсального договора в цифровой среде // Системы и средства информатики, 2020. Т. 30. № 3. С. 154–162.
15. *Ilyin A. V., Ilyin V. D.* The technologies of commodity-money circulation on the basis of personal and corporative e-banks // Int. J. Open Information Technologies, 2020. Vol. 8. Iss. 6. P. 81–84.
16. Куроев Б. Н. Сравнение эффективности алгоритмов управления с учетом точности данных и реализации решений // Управление большими системами, 2011. Вып. 34. С. 279–291.
17. Ильин В. Д., Гавриленко Ю. В., Ильин А. В., Макаров Е. М. Математические средства ситуационной информатизации. — М.: Наука, Физматлит, 1996. 88 с.
18. The Digital Twin // General Electric, 2018. https://www.ge.com/digital/sites/default/files/The-Digital-Twin_Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf.
19. *Ilyin A. V., Ilyin V. D.* Variational online budgeting taking into account the priorities of expense items // Agris On-line Papers in Economics and Informatics, 2016. Vol. 8. Iss. 3. P. 51–56.
20. *Ilyin A. V., Ilyin V. D.* Solving situationally definable linear problems of resource planning: A review of updated technology // Информационные технологии и вычислительные системы, 2019. № 3. Р. 99–106.
21. Ильин А. В. Экспертное планирование ресурсов. — М.: ИПИ РАН, 2013. 58 с.
22. Ильин В. Д. Технология назначенных платежей в среде цифровых двойников // Системы и средства информатики, 2018. Т. 28. № 3. С. 227–235. doi: 10.14357 / 08696527180318.
23. *Ilyin A. V., Ilyin V. D.* E-trade with direct lending and normalized money // Agris on-line Papers in Economics and Informatics, 2015. Vol. 7. Iss. 4. P. 57–64.

Поступила в редакцию 14.08.20

BLOCKCHAIN TECHNOLOGIES FOR NORMALIZED BUDGET SUPPORT OF NATIONAL PROJECTS

A. V. Ilyin¹ and V. D. Ilyin²

¹State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation

²A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

Abstract: Blockchain technologies for normalized budget support (NBS-technologies) are considered as means necessary for resource support of design and implementation of national projects for protection and development of the country's potential. The part of the methodological support for development of NBS-technologies considered in the article includes methodologies of situational online budgeting and normalized commodity-money circulation. Online budget planning is considered as a problem of the interval cost planning, taking into account the situationally dependent mandatory and orienting requirements for the desired solution. The problem is solved by means of the method of target displacement of solution in the computational experiment mode. The methodology of normalized commodity-money circulation includes methods for formation and implementation of contractual relations in the digital environment, payment and commodity lending, and online banking based on banks-providers, corporate, and personal electronic banks. Blockchain technologies for normalized budget support practically eliminate the unplanned use of budget funds.

Keywords: blockchain technologies for normalized budget support (NBS-technologies); resource support of national projects for protection and development of the country's potential; situational online budgeting; normalized commodity-money circulation

DOI: 10.14357/08696527200414

References

1. Natsional'nye proekty: tselevye pokazateli i osnovnye rezul'taty [National projects: Target indicators and main results]. 2019. Moscow. Available at: <http://static.government.ru/media/files/p7nn2CS0pVhvQ98OOwAt2dzClAietQih.pdf> (accessed August 5, 2020).
2. Ilyin, V. D. 1996. *Osnovaniya situatsionnoy informatizatsii* [Fundamentals of situational informatization]. Moscow: Nauka, Fizmatlit. 180 p.
3. Tapscott, D. 1996. *The digital economy: Promise and peril in the age of networked intelligence*. New York, NY: McGraw-Hill. 342 p.
4. Christensen, C. M. 1997. *The innovator's dilemma: When new technologies cause great firms to fail*. Boston, MA: Harvard Business School Press. 288 p.

5. Ilyin, A. V., and V. D. Ilyin. 2014. Towards a normalized economic mechanism based on E-services. *Agris on-line Papers in Economics and Informatics* 6(3):39–49
6. Oxford Economics. 2015. The new digital economy: How it will transform business. Available at: <http://www.pwc.com/mt/en/publications/assets/the-new-digital-economy.pdf> (accessed August 5, 2020).
7. G20 digital economy development and cooperation initiative. 2016. G20 Summit. Available at: <http://en.kremlin.ru/supplement/5111> (accessed July 5, 2020).
8. Tsifrovaya ekonomika Rossiyskoy Federatsii: Programma, utverzhдennaya rasporyazheniem Pravitel'stva Rossiyskoy Federatsii [Digital economy of the Russian Federation: Program approved by Order No. 1632-r dated July 28, 2017 of the Government of the Russian Federation]. Available at: <http://d-russia.ru/wpcontent/uploads/2017/07/programma-tsifrov-econ.pdf> (accessed July 5, 2020).
9. Ilyin, A. V., and V. D. Ilyin. 2019. The normalized economic mechanism in the digital environment. *Int. J. Open Information Technologies* 7(12):77–83.
10. Ilyin, A. V., and V. D. Ilyin. 2020. Normalizovanny ekonomicheskiy mekhanizm: tsifrovye tekhnologii polivalyutnogo rynka [The normalized economic mechanism: Digital technologies of the multicurrency market]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(1):186–197. doi: 10.14357/08696527200116.
11. Kompaniya MINDSMITH provela masshtabnoe issledovanie otechestvennogo blokcheyn-rynska [MINDSMITH company performed a large-scale study of the Russian blockchain market]. Available at: <http://nbj.ru/pubs/banki-i-biznes/2019/11/26/kompanija-mindsmith-provela-masshtabnoe-issledovanie-otechestvennogo-blokchein-rynska/?full> (accessed August 5, 2020).
12. Ilyin, V. D. 2018. Simvol'noe modelirovanie [Symbolic modeling]. *Bol'shaya rossiyskaya entsiklopediya* [The Great Russian Encyclopedia]. Available at: <http://dev.bigenc.ru/technology and technique/text/4010980> (accessed July 5, 2020).
13. Ilyin, A. V., and V. D. Ilyin. 2017. Osnovy kontseptsii situatsionnoy informatizatsii zhiznedeyatel'nosti [Basics of the concept of situational informatization of population activities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(3):188–201. doi: 10.14357/08696527170315.
14. Ilyin, A. V., and V. D. Ilyin. 2020. Situationsnaya model' universal'nogo dogovora v tsifrovoy srede [Situational model of universal agreement in a digital environment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(3):154–162.
15. Ilyin, A. V., and V. D. Ilyin. 2020. The technologies of commodity-money circulation on the basis of personal and corporative e-banks. *Int. J. Open Information Technologies* 8(6):81–84.
16. Kurov, B. N. 2011. Sravnenie effektivnosti algoritmov upravleniya s uchetom tochnosti dannykh i realizatsii resheniy [Comparison of the effectiveness of control algorithms taking into account the accuracy of data and implementation of solutions]. *Upravlenie bol'shimi sistemami* [Large-Scale Systems Control] 34:279–291.
17. Ilyin, V. D., Yu. V. Gavrilenko, A. V. Ilyin, and E. M. Makarov. 1996. *Matematicheskie sredstva situatsionnoy informatizatsii* [Mathematical means of situational informatization]. Moscow: Nauka, Fizmatlit. 88 p.
18. General Electric. 2018. The Digital Twin. Available at: <https://www.ge.com/digital/sites/default/files/The-Digital-Twin.Compressing-Time-to-Value-for-Digital-Industrial-Companies.pdf> (accessed August 5, 2020).

19. Ilyin, A. V., and V. D. Ilyin. 2016. Variational online budgeting taking into account the priorities of expense items. *Agris on-line Papers in Economics and Informatics* 8(3):51–56.
20. Ilyin, A. V., and V. D. Ilyin. 2019. Solving situationally definable linear problems of resource planning: A review of updated technology. *Informatsionnye tekhnologii i vychislitel'nye sistemy* [J. Information Technologies Computing Systems] 3:99–106.
21. Ilyin, A. V. 2013. *Ekspertnoe planirovanie resursov* [Expert resource planning]. Moscow: IPI RAN. 58 p.
22. Ilyin, V. D. 2018. Tekhnologiya naznachennykh platezhey v srede tsifrovых dvoynikov [Designated payments technology in digital twins environment]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 28(3):227–235. doi: 10.14357/08696527180318.
23. Ilyin, A. V., and V. D. Ilyin. 2015. E-trade with direct lending and normalized money. *Agris on-line Papers in Economics and Informatics* 7(4):57–64.

Received August 14, 2020

Contributors

Ilyin Alexander V. (b. 1975) — Candidate of Science (PhD) in technology, leading scientist, State Research Institute of Aviation Systems, 7 Viktorenko Str., Moscow 125319, Russian Federation; ilyin@res-plan.com

Ilyin Vladimir D. (b. 1937) — Doctor of Science in technology, professor, leading scientist, A. A. Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; vdilyin@yandex.ru

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБРАБОТКИ ИНФОРМАЦИИ НЕЙРОННЫМИ СЕТЯМИ С ИСПОЛЬЗОВАНИЕМ ПАРАБОЛИЧЕСКИХ ИНТЕГРОДИФФЕРЕНЦИАЛЬНЫХ СПЛАЙНОВ В КАЧЕСТВЕ ФУНКЦИЙ АКТИВАЦИИ НЕЙРОНОВ

Т. К. Бирюкова¹

Аннотация: Рассмотрены способы повышения эффективности обработки информации нейронными сетями (нейросетями) путем использования разработанных автором параболических интегродифференциальных сплайнов (ИД-сплайнов) как функций активации нейронов. Предложено сделать коэффициенты параболических ИД-сплайнов обучаемыми параметрами нейросети наряду с весами нейронов. В этом случае функция активации в виде параболического ИД-сплайна изменяется в процессе обучения так, чтобы минимизировать функцию ошибки, за счет чего повышается точность результатов работы нейросети, сокращается время ее обучения и работы в режиме эксплуатации. Проанализированы перспективы модификации нейросетей с известными архитектурами (таких как ResNet) путем замены в них функций активации на ИД-сплайновые функции активации. Представляется, что такой подход сможет повысить качество функционирования ряда популярных нейросетей. Сделан вывод о том, что параболические ИД-сплайны в качестве функций активации могут повысить эффективность технологий искусственного интеллекта, в частности в таких задачах, как принятие решений, создание компьютерных игр, аппроксимация и прогнозирование данных (в финансовой и социальной сфере, в науке и др.), классификация информации, обработка изображений и видеороликов, применение компьютерного зрения, обработка текстов, речи, музыки и др.

Ключевые слова: искусственный интеллект; машинное обучение; глубокое обучение; нейронная сеть; нейросеть; функция активации; сплайновая функция активации; интегродифференциальный сплайн; ИД-сплайн; параболический сплайн

DOI: 10.14357/08696527200415

1 Введение

В настоящее время методы искусственного интеллекта, основанные на применении нейросетей, становятся одним из основных мировых трендов развития

¹Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, yukonta@mail.ru

информационных технологий [1]. Ключевую роль в работе нейросетей играют функции активации, которые определяют выходные сигналы нейронов. В данной работе рассмотрены способы повышения эффективности обработки информации нейросетями путем использования разработанных автором параболических ИД-сплайнов [2] как функций активации нейронов.

Коэффициенты параболических ИД-сплайнов

$$S_{2\text{ИД}}(x) = \bigcup_{i=0}^{n-1} S_{2\text{ИД},i}(x) \quad (1)$$

вычисляются с использованием условия нулевой интегральной невязки между составляющими сплайнами (1) многочленами $S_{2\text{ИД},i}(x)$ и соответствующими фрагментами аппроксимируемой функции $f(x)$ на отрезках $[\bar{x}_i, \bar{x}_{i+1}]$, ограниченных узлами сетки сплайна θ : $\bar{x}_0 < \dots < \bar{x}_i < \dots < \bar{x}_n$:

$$\int_{\bar{x}_i}^{\bar{x}_{i+1}} S_{2\text{ИД},i}(x) dx - I_i^{i+1} = 0, \quad (2)$$

где $I_i^{i+1} \approx \int_{\bar{x}_i}^{\bar{x}_{i+1}} f(x) dx$ — известные или вычисленные по каким-либо квадратурным формулам значения интеграла от аппроксимируемой функции $f(x)$ на отрезке $[\bar{x}_i, \bar{x}_{i+1}]$.

Если обозначить $u = (x - \bar{x}_i)/h_{i+1}$, то звенья $S_{2\text{ИД},i}(x)$ сплайна (1) будут иметь вид:

$$S_{2\text{ИД},i}(u) = (-6u^2 + 6u) \frac{I_i^{i+1}}{h_{i+1}} + (3u^2 - 4u + 1) f_i + (3u^2 - 2u) f_{i+1}, \quad (3)$$

где $h_{i+1} = \bar{x}_{i+1} - \bar{x}_i$; f_i — известные (возможно, с некоторой погрешностью) или вычисленные значения функции $f(x)$ в узлах сетки θ ($f_i \approx f(\bar{x}_i)$).

Коэффициенты многочленов (3) определяются интегральным условием согласования (2) и двумя функциональными условиями согласования на концах отрезка $[\bar{x}_i, \bar{x}_{i+1}]$: $S_{2\text{ИД},i}(\bar{x}_i) = f_i$ и $S_{2\text{ИД},i}(\bar{x}_{i+1}) = f_{i+1}$.

При вычислении параметров f_i ($i = 0, \dots, n$) из следующей системы линейных алгебраических уравнений (СЛАУ) с граничными условиями, определяемыми из физического смысла решаемой задачи, первая производная сплайна всегда непрерывна, т. е. сплайн имеет дефект 1:

$$\frac{1}{h_i} f_{i-1} + 2 \left(\frac{1}{h_i} + \frac{1}{h_{i+1}} \right) f_i + \frac{1}{h_{i+1}} f_{i+1} = 3 \left(\frac{I_i^{i+1}}{h_{i+1}^2} + \frac{I_{i-1}^i}{h_i^2} \right), \\ i = 1, \dots, n-1. \quad (4)$$

В качестве граничных уравнений, дополняющих СЛАУ (4), можно использовать, например, значения сеточной функции в точках \bar{x}_0 и \bar{x}_n :

$$f_0 = f(\bar{x}_0), \quad f_n = f(\bar{x}_n). \quad (5)$$

Система (4) в сочетании с уравнениями (5) — это трехдиагональная СЛАУ с диагональным преобладанием (что обеспечивает устойчивость вычислений), решаемая методом прогонки [2].

Относительно малое количество вычислительных операций в формуле (3) по сравнению с традиционными параболическими (строящимися со сдвигом узлов сплайна относительно точек аппроксимируемой сеточной функции) и кубическими сплайнами [3] позволило предположить, что параболический ИД-сплайн имеет широкие перспективы по применению в нейросетях в качестве адаптивной (изменяющейся при обучении) функции активации.

В 1996 г. была опубликована статья [4], описывающая функции активации в виде кубических сплайнов. Указано, что их преимущество перед неизменяемыми функциями активации заключается в том, что сплайновые функции активации настраиваются в процессе обучения нейросети для оптимизации ее работы. В последующих статьях по данному вопросу рассматривались различные способы построения архитектуры нейросетей с такими функциями активации и примеры их использования в нейросетях [5, 6]. Однако упомянутые в этих статьях кубические сплайны требуют больше вычислительных ресурсов, чем параболические ИД-сплайны со звеньями (3). И поскольку нейросети обучаются на больших объемах данных (обучение иногда занимает месяцы), а функция активации вычисляется некоторое число раз (определенное архитектурой сети) для каждого сигнала, обрабатываемого сетью, то сокращение времени вычисления значений функции активации играет значительную роль.

Таким образом, применение параболических ИД-сплайнов в нейросетях со сложной архитектурой позволит сократить время обучения сети и ее работы в режиме эксплуатации по сравнению с кубическими сплайнами. В то же время точность результатов, выдаваемых нейросетью с ИД-сплайновой функцией активации, будет выше, чем при неизменяемых функциях активации. В частности, перспективной выглядит идея по замене неадаптируемых функций активации на параболические ИД-сплайны в популярных в настоящее время нейросетях — например, в сети ResNet, применяемой в основном для анализа изображений и превосходящей человеческий уровень классификации изображений.

2 Обработка информации нейронными сетями с использованием параболических интегродифференциальных сплайнов как функций активации

В данном разделе предлагается алгоритм применения параболического ИД-сплайна в качестве функции активации. Для построения алгоритма рассмотрим вначале принцип работы одного нейрона в нейросети.

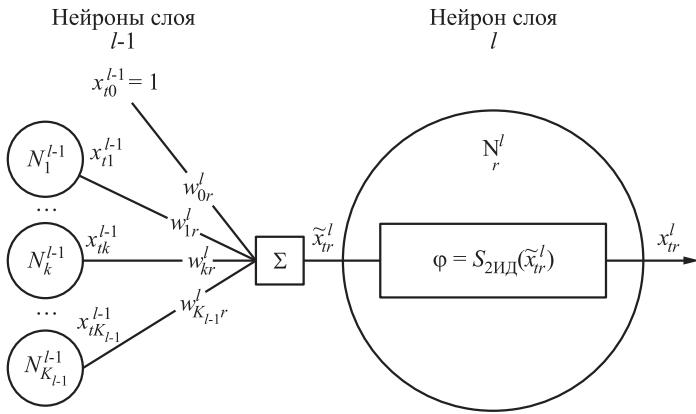


Схема работы одного нейрона в «полносвязном» слое нейросети с параболическим ИД-сплайном в качестве функции активации

Искусственная нейронная сеть строится по принципам работы человеческого мозга. Каждый нейрон в нейросети — это объект, преобразующий сигналы, полученные от некоторого множества нейронов, и отправляющий другим нейронам сигналы, являющиеся результатами этого преобразования. Нейроны объединяются в группы — слои. Каждый слой обрабатывает сигналы по определенному правилу. Множество сигналов, подаваемых в некоторый момент времени на вход нейросети, представляет собой массив числовых величин (в общем случае многомерный), например: характеристики некоторого объекта, оцифрованное изображение, текст, звуковая дорожка. Такой массив в технической литературе, посвященной нейросетевым технологиям, обозначается термином «тензор» (здесь нет связи с объектом «тензор» в линейной алгебре).

Нейросеть содержит изменяемые параметры, которые настраиваются так, чтобы сеть оптимальным образом выполняла поставленную задачу. Для решения задач различных типов используются нейросети, состоящие из слоев, построенных по разным принципам (отличающихся числом нейронов и способом преобразования сигналов).

В качестве примера на рисунке показана схема работы одного нейрона N_r^l так называемого «полносвязного» слоя нейросети. Полносвязный слой характерен тем, что в нем каждый нейрон получает сигналы от всех нейронов предыдущего слоя и выдает сигналы всем нейронам следующего слоя.

На выходе из $(l - 1)$ -го слоя, состоящего из K_{l-1} нейронов (см. рисунок), получаются векторы

$$X_t^{l-1} = \left(x_{t0}^{l-1} \dots x_{tk}^{l-1} \dots x_{tK_{l-1}}^{l-1} \right), \quad t = 1, \dots, T,$$

длиной K_{l-1} .

Каждый r -й нейрон \mathcal{N}_r^l l -го («полносвязного») слоя последовательно в моменты времени t ($t = 1, \dots, T$) принимает векторы

$$X_t^{l-1} = (x_{t1}^{l-1} \dots x_{tk}^{l-1} \dots x_{tK_{l-1}}^{l-1})$$

от нейронов предыдущего, $(l-1)$ -го слоя, при этом каждый сигнал x_{tk}^{l-1} умножается на некоторый вес w_{kr}^l , характеризующий связь между нейронами \mathcal{N}_k^{l-1} и \mathcal{N}_r^l . Далее нейрон \mathcal{N}_r^l выполняет линейное преобразование входных сигналов:

$$\tilde{x}_{tr}^l = \sum_{k=0}^{K_{l-1}} x_{tk}^{l-1} w_{kr}^l \quad (\text{здесь } x_{t0}^{l-1} = 1), \quad (6)$$

где K_{l-1} — число нейронов $(l-1)$ -го слоя, от которых нейрон \mathcal{N}_r^l получает сигналы. Веса w_{kr}^l ($k = 1, \dots, K_{l-1}$) связей нейрона \mathcal{N}_r^l с каждым из взаимодействующих с ним нейронов предыдущего, $(l-1)$ -го слоя составляют множество весов нейрона \mathcal{N}_r^l . Дополнительно в число весов нейрона включается величина w_{0r}^l , называемая «смещением» (и x_{t0}^{l-1} в линейной комбинации (6) принимается равным 1).

Веса нейросети, включая смещения, являются обучаемыми параметрами. Обучение состоит в корректировке весов так, чтобы улучшить качество работы нейросети. В зависимости от задач, решаемых нейросетью, используются разные способы обучения. При так называемом «обучении с учителем» предсказанные нейросетью результаты сравниваются с известными значениями («правильными ответами») путем вычисления значения некоторой функции ошибки (называемой также функцией потерь). Далее веса нейросети корректируются так, чтобы минимизировать функцию потерь. Обычно для этого используется метод градиентного спуска и его модификации. Наиболее распространенной функцией потерь служит среднеквадратичная ошибка. Обучение проводится итерациями. Тренировочный набор данных рекомендуется делить на пакеты (так называемые батчи). В одной итерации на вход нейросети подается один батч тензоров тренировочных данных, и после обработки батча всеми слоями обновляются обучаемые параметры сети.

Выходной сигнал нейрона \mathcal{N}_k^P — это результат применения к сумме (6) нелинейной функции φ , называемой **функцией активации**: $x_{tr}^l = \varphi(\tilde{x}_{tr}^l)$.

К наиболее часто используемым функциям активации относятся: сигмоида, гиперболический тангенс, ReLU (и ее модификации), ELU [5]. В разных слоях нейросети могут применяться разные функции активации, но в одном слое обычно используется одна и та же функция активации. Вышеуказанные функции активации не изменяются при обучении нейросети.

В данной работе предлагается в качестве функции активации использовать параболический ИД-сплайн со звеньями (3), где интегралы I_i^{i+1} являются обучаемыми параметрами нейросети наряду с весами нейронов. Форма кривой такой

функции активации изменяется в процессе обучения так, чтобы улучшить результаты работы нейросети.

В «полносвязных» слоях с малым числом нейронов (менее 50) можно для каждого нейрона использовать свою ИД-сплайновую функцию активации так, чтобы интегральные параметры I_i^{i+1} , входящие в состав коэффициентов сплайна, менялись независимо для разных нейронов. Но в более сложных нейросетях (например, со «сверточными» слоями, используемыми, в частности, для обработки изображений) построение большого числа разных сплайнов вызвало бы неприемлемое увеличение времени обучения. Поэтому предлагается в больших нейросетях использовать в одном слое одну ИД-сплайновую функцию активации (но в разных слоях — разные, меняющиеся независимо друг от друга).

Ниже приводится алгоритм вычисления значений ИД-сплайновой функции от сигналов, выдаваемых слоем нейронов.

Рассмотрим работу l -го слоя нейросети, после которого вызывается функция активации. l -й слой нейросети выдает батч $B_{\tilde{X}}$, состоящий в общем случае из многомерных тензоров \tilde{X}_t ($t = 1, \dots, T$), поступающий на вход функции активации: $B_{\tilde{X}} = (\tilde{X}_1, \dots, \tilde{X}_t, \dots, \tilde{X}_T)^T$. Возьмем сетку узлов сплайна θ . Для вычисления значений элементов тензора выходных сигналов ИД-сплайновой функции активации $B_Y = (Y_1, \dots, Y_t, \dots, Y_T)^T$ нужно для каждого элемента \tilde{x} каждого тензора \tilde{X}_t из батча $B_{\tilde{X}}$ определить интервал сетки узлов $[\bar{x}_i, \bar{x}_{i+1}]$, куда «попал» \tilde{x} : $\tilde{x} \in [\bar{x}_i, \bar{x}_{i+1}]$. Это необходимо для того, чтобы понять, какое звено сплайна (т. е. какие параметры I_i^{i+1}, f_i, f_{i+1}) нужно использовать для вычисления выходного значения $y = S_{2\text{ИД}, i}(\tilde{x})$, где i определяется условием $\tilde{x} \in [\bar{x}_i, \bar{x}_{i+1}]$. Для каждого батча $B_{\tilde{X}}$ нужно составить следующие тензоры (имеющие такой же размер, как $B_{\tilde{X}}$): B_{ind} и $B_{\text{ind}1}$, где каждый элемент представляет собой индекс i и $i + 1$ (соответственно) левой и правой граничной точки интервала сетки θ , куда «попал» соответствующий элемент \tilde{x} батча $B_{\tilde{X}}$ ($\tilde{x} \in [\bar{x}_i, \bar{x}_{i+1}]$); B_{IH} — состоящий из значений I_i^{i+1}/h_{i+1} ; B_f и B_{f1} — состоящие из значений f_i и f_{i+1} соответственно; B_u — состоящий из значений $(\tilde{x} - \bar{x}_i)/h_{i+1}$.

Далее по формуле, вытекающей из (3), можно вычислить:

$$\begin{aligned} B_Y &= S_{2\text{ИД}}(B_{\tilde{X}}) = \\ &= (-6B_u^2 + 6B_u) B_{IH} + (3B_u^2 - 4B_u + 1) B_f + (3B_u^2 - 2B_u) B_{f1}. \end{aligned}$$

(Здесь операции умножения и сложения — это почленное умножение и сложение значений многомерных массивов (тензоров) соответственно.)

На каждой итерации обучения (после обработки нейросетью одного батча из тренировочного набора данных) интегральные параметры I_i^{i+1} ($i = 0, \dots, n - 1$) изменяются. После обновления I_i^{i+1} ($i = 0, \dots, n - 1$) нужно вычислить значения f_i ($i = 0, \dots, n$) из СЛАУ (4) в сочетании с (5).

Перед началом обучения следует инициализировать параметры I_i^{i+1} ($i = 0, \dots, n - 1$). Их можно задать как интегралы от известной функции активации

ции (из перечисленных выше), используемой для задач, аналогичных решаемой. Вычислить I_i^{i+1} можно как по явным формулам (пронтегрировав эти известные функции), так и с помощью квадратурных формул вычислительной математики [2].

3 Заключение

Функция активации в виде параболического ИД-сплайна меняется в процессе обучения так, чтобы минимизировать функцию ошибки, за счет чего происходит повышение точности результатов работы нейросети, сокращение времени ее обучения и ускорение работы в режиме эксплуатации. В этом случае применение нейросетей с упрощенной архитектурой (с меньшим числом слоев, нейронов) не вызовет потерю качества.

Представляется перспективным направлением модификация нейросетей с известными архитектурами (таких как ResNet) путем замены в них функций активации на ИД-сплайновые функции активации. При этом остается возможность использования предобученных весов нейронов, поставляемых стандартными программными библиотеками для создания нейросетей (например, PyTorch), что очень важно для сокращения времени обучения.

Возможность применения как равномерных, так и неравномерных сеток узлов параболического ИД-сплайна позволяет повысить точность результатов работы нейросети. Скорость обучения и работы в режиме эксплуатации ИД-сплайновой нейросети не зависит от вида сетки сплайна.

Параболические ИД-сплайны в качестве функций активации могут повысить эффективность технологий искусственного интеллекта в таких задачах, как принятие решений, создание компьютерных игр, аппроксимация и прогнозирование данных (в финансовой и социальной сфере, в науке и др.), классификация информации, обработка изображений и видеороликов, применение компьютерного зрения, обработка текстов, речи, музыки и др.

Литература

1. Зацаринный А. А., Киселев Э. В., Козлов С. В., Колин К. К. Информационное пространство цифровой экономики России. Концептуальные основы и проблемы формирования. — М: ФИЦ ИУ РАН, 2018. 236 с.
2. Киреев В. И., Бирюкова Т. К. Интегродифференциальный метод обработки информации и его применение в численном анализе. — М.: ИПИ РАН, 2014. 267 с.
3. Стечкин С. Б., Субботин Ю. Н. Сплайны в вычислительной математике. — М.: Наука, 1976. 248 с.
4. Campolucci P., Cappellotti F., Guarneri S., Piazza F., Uncini A. Neural networks with adaptive spline activation function // 8th Mediterranean Electrotechnical Conference on Industrial Applications in Power Systems, Computer Science and Telecommunications Proceedings. — IEEE, 1996. Р. 1442–1445.

5. Mayer H. A., Schwaiger R. Evolution of cubic spline activation functions for artificial neural networks // Progress in artificial intelligence, knowledge extraction, multi-agent systems, logic programming and constraint solving / Eds. P. Brazdil, A. Jorge. — Lecture notes in computer science ser. — Springer, 2001. Vol. 2258. P. 63–73.
6. Scardapane S., Scarpiniti M., Comminiello D., Uncini A. Neural advances in processing nonlinear dynamics signals: Learning activation functions from data using cubic spline interpolation // Neural advances in processing nonlinear dynamic signals / Eds. A. Esposito, M. Faundez-Zanuy, F. Morabito, E. Pasero. — Smart innovation, systems and technologies ser. — Springer, 2017. Vol. 102. P. 73–83.

Поступила в редакцию 02.11.20

PARABOLIC INTEGRODIFFERENTIAL SPLINES AS ACTIVATION FUNCTIONS TO INCREASE THE EFFICIENCY OF INFORMATION PROCESSING BY NEURAL NETWORKS

T. K. Biryukova

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

Abstract: The paper considers the method to increase the efficiency of information processing by neural networks by using the parabolic integrodifferential splines (ID-splines) developed by the author as an activation function (AF) for neurons. If the coefficients of parabolic ID-splines along with the weights of the neurons are the trainable parameters of the neural network, then the AF in the form of a parabolic ID spline changes in the learning process to minimize the error function. This increases the accuracy of the results of the neural network calculations and accelerates its training and operation. The prospects for modifying neural networks with known architectures (such as ResNet) by introducing ID-spline as AF are analyzed. Apparently, such an approach can improve the quality of functioning of some popular neural networks. It is concluded that parabolic ID splines as AF can increase the efficiency of artificial intelligence technologies in such tasks as decision making, computer games development, approximating and predicting data (in the financial and social spheres, in science, etc.), classification of information, processing of images and videos, application of computer vision, processing of texts, speech, and music, etc.

Keywords: artificial intelligence; deep learning; neural network; activation function; spline interpolation; integrodifferential spline; parabolic spline

DOI: 10.14357/08696527200415

References

1. Zatsarinnyy, A. A., E. V. Kiselev, S. V. Kozlov, and K. K. Kolin. 2018. *Informatsionnoe prostranstvo tsifrovoy ekonomiki Rossii. Kontseptual'nye osnovy i problemy formirovaniya* [Information space of the digital economy of Russia. Conceptual foundations and problems of formation]. Moscow: FRC CSC RAS. 236 p.
2. Kireev, V. I., and T. K. Biryukova. 2014. *Integrodifferentsial'nyy metod obrabotki informatsii i ego primenenie v chislennom analize* [Integrodifferential method of information processing and its application in numerical analysis]. Moscow: IPI RAN. 267 p.
3. Stechkin, S. B., and Yu. N. Subbotin. 1976. *Splayny v vychislitel'noy matematike* [Splines in computational mathematics]. Moscow: Nauka. 248 p.
4. Campolucci, P., F. Cappellari, S. Guarnieri, F. Piazza, and A. Uncini. 1996. Neural networks with adaptive spline activation function. *8th Mediterranean Electrotechnical Conference on Industrial Applications in Power Systems, Computer Science and Telecommunications Proceedings*. IEEE. 1442–1445.
5. Mayer, H. A., and R. Schwaiger. 2001. Evolution of cubic spline activation functions for artificial neural networks. *Progress in artificial intelligence, knowledge extraction, multi-agent systems, logic programming and constraint solving*. Eds. P. Brazdil and A. Jorge. Lecture notes in computer science ser. Springer. 2258:63–73.
6. Scardapane, S., M. Scarpiniti, D. Comminiello, and A. Uncini. 2017. Learning activation functions from data using cubic spline interpolation. *Neural advances in processing nonlinear dynamics signals: Learning activation functions from data using cubic spline interpolation*. Eds. A. Esposito, M. Faundez-Zanuy, F. Morabito, and E. Pasero. Smart innovation, systems and technologies ser. Springer. 102:73–83.

Received November 2, 2020

Contributor

Biryukova Tatiana K. (b. 1968) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; yukonta@mail.ru

О Б А В Т О Р АХ

Адамович Игорь Михайлович (р. 1934) — кандидат технических наук, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Бирюкова Татьяна Константиновна (р. 1968) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Волков Олег Игоревич (р. 1964) — ведущий программист Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Гайдамака Юлия Васильевна (р. 1971) — доктор физико-математических наук, профессор кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Грушо Александр Александрович (р. 1946) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Дьяченко Юрий Георгиевич (р. 1958) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Егорова Анна Юрьевна (р. 1991) — младший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ерофеева Арина Владимировна (р. 1997) — студентка кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Захарова Татьяна Валерьевна (р. 1962) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Зацман Игорь Моисеевич (р. 1952) — доктор технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Ильин Александр Владимирович (р. 1975) — кандидат технических наук, ведущий научный сотрудник Государственного научно-исследовательского института авиационных систем

Ильин Владимир Дмитриевич (р. 1937) — доктор технических наук, профессор, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кончеков Евгений Михайлович (р. 1988) — кандидат физико-математических наук, старший научный сотрудник отдела физики плазмы Института общей физики им. А. М. Прохорова Российской академии наук; доцент кафедры физики Российского национального исследовательского медицинского университета им. Н. И. Пирогова

Кривенко Михаил Петрович (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кружков Михаил Григорьевич (р. 1975) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Кудрявцев Алексей Андреевич (р. 1978) — кандидат физико-математических наук, доцент кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Ларичев Владимир Андреевич (р. 1972) — кандидат физико-математических наук, научный сотрудник отдела геофизической акустики Акустического института им. академика Н. Н. Андреева

Листопад Сергей Викторович (р. 1984) — кандидат технических наук, старший научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Максимов Герман Адольфович (р. 1963) — доктор физико-математических наук, начальник отдела геофизической акустики Акустического института им. академика Н. Н. Андреева

Маньяков Юрий Анатольевич (р. 1984) — старший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Медведева Екатерина Георгиевна (р. 1985) — ассистент кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Морозов Николай Викторович (р. 1956) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Николаев Андрей Владимирович (р. 1973) — кандидат физико-математических наук, научный сотрудник Федерального исследовательского центра химической физики им. Н. Н. Семёнова Российской академии наук

Нуриев Виталий Александрович (р. 1980) — кандидат филологических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Писковский Виктор Олегович (р. 1963) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Поляков Никита Алексеевич (р. 1997) — студент магистратуры кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Рацеев Сергей Михайлович (р. 1979) — доктор физико-математических наук, доцент кафедры информационной безопасности и теории управления Ульяновского государственного университета

Румовская София Борисовна (р. 1985) — кандидат технических наук, научный сотрудник Калининградского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук; старший преподаватель Балтийского федерального университета имени Иммануила Канта

Сенчило Владимир Викторович (р. 1963) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Скворцова Нина Николаевна (р. 1956) — доктор физико-математических наук, профессор кафедры моделирования радиофизических процессов МИРЭА — Российского технологического университета; профессор кафедры № 69 Национального исследовательского ядерного университета «МИФИ»; ведущий научный сотрудник отдела физики плазмы Института общей физики им. А. М. Прохорова Российской академии наук

Смирнов Виталий Александрович (р. 1990) — аспирант кафедры моделирования радиофизических процессов МИРЭА — Российского технологического университета; ведущий инженер отдела гидроакустических преобразователей Акустического института им. академика Н. Н. Андреева

Сорокин Андрей Иванович (р. 1987) — младший научный сотрудник Орловского филиала Федерального исследовательского центра «Информатика и управление» Российской академии наук

Степченков Дмитрий Юрьевич (р. 1973) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Степченков Юрий Афанасьевич (р. 1951) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Терёхина Ирина Юрьевна (р. 1994) — аспирант факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Тимонина Елена Евгеньевна (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Хайров Эмиль Маратович (р. 1997) — студент магистратуры кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов

Хилько Дмитрий Владимирович (р. 1987) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Череватенко Ольга Ивановна (р. 1983) — кандидат физико-математических наук, доцент кафедры высшей математики Ульяновского государственного педагогического университета имени И. Н. Ульянова

Шестаков Олег Владимирович (р. 1976) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

Шоргин Сергей Яковлевич (р. 1952) — доктор физико-математических наук, профессор, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2020 г.

	№	Стр.
Адамович И. М., Волков О. И. Автоматизированная оценка достоверности конкретно-исторических фактов	1	160–170
Адамович И. М., Волков О. И. Автоматизированный поиск противоречий в конкретно-исторической информации	3	145–153
Адамович И. М., Волков О. И. Механизм темпорального со- поставления конкретно-исторических фактов	4	138–146
Адамович И. М., Волков О. И. Модель функционирования системы автоматизированной оценки достоверности конкрет- но-исторических фактов	2	177–186
Бирюкова Т. К. Повышение эффективности обработки инфор- мации нейронными сетями с использованием параболических интегродифференциальных сплайнов в качестве функций активации нейронов	4	159–167
Борисов А. В., Босов А. В., Жуков Д. В., Иванов А. В., Сушко Д. В. Информационные аспекты обеспечения без- опасности на транспорте: онтология предметной области, модели и варианты использования	1	126–134
Босов А. В., Сучков А. П. Информационно-аналитический си- туационный центр стратегического планирования в области обеспечения национальной безопасности: факторы влияния	1	115–125
Босов А. В., Сучков А. П. Информационно-аналитический си- туационный центр стратегического планирования в области обеспечения национальной безопасности: перспективные архитектурные решения	2	145–162
Босов А. В. см. Борисов А. В.		
Вакуленко В. В., Гончаров А. А., Дурново А. А., Зац- ман И. М. Задачи базы данных фразеологического словаря и стадии ее проектирования	2	113–123
Водолаженко А. А. см. Кабанов А. С.		
Волков О. И. см. Адамович И. М.		
Волков О. И. см. Адамович И. М.		
Волков О. И. см. Адамович И. М.		
Волков О. И. см. Адамович И. М.		
Гаврилов В. Е., Зацаринный А. А. Актуальные проблемы ин- формационной безопасности в условиях пандемии корона- вируса	3	123–132
Гайдамака Ю. В. см. Медведева Е. Г.		
Гаранин А. И. см. Зацаринный А. А.		
Гончаров А. А. см. Вакуленко В. В.		

	№	Стр.
Гринченко С. Н. Базисные информационные технологии и «промышлённые революции»: какова их взаимосвязь?	3	163–171
Грушо А. А., Грушо Н. А., Николаев А. В., Писковский В. О., Сенчило В. В., Тимонина Е. Е. Порождение угроз в гетерогенных системах цифровой экономики	1	46–55
Грушо А. А., Грушо Н. А., Сенчило В. В., Тимонина Е. Е. Предотвращение массового вывода из строя слабозащищенных мобильных систем информационной поддержки цифровой экономики	2	21–30
Грушо А. А., Зацаринный А. А., Тимонина Е. Е. Описание динамики цифровой экономики с помощью электронной бухгалтерской книги	1	108–114
Грушо А. А., Николаев А. В., Писковский В. О., Сенчило В. В., Тимонина Е. Е. Комплексное обеспечение информационной безопасности частных облачных вычислительных сред	4	76–82
Грушо А. А. см. Терёхина И. Ю.		
Грушо Н. А. см. Грушо А. А.		
Грушо Н. А. см. Грушо А. А.		
Дулин С. К., Никишин Д. А. Подходы к интеграции прикладных концептуальных схем в составе унифицированной геоонтологии	2	68–77
Дурново А. А. см. Вакуленко В. В.		
Дьяченко Д. Ю. см. Степченков Ю. А.		
Дьяченко Д. Ю. см. Степченков Ю. А.		
Дьяченко Ю. Г. см. Степченков Ю. А.		
Дьяченко Ю. Г. см. Степченков Ю. А.		
Дьяченко Ю. Г. см. Степченков Ю. А.		
Егоров В. Б. Некоторые вопросы программного определения центров обработки данных	2	103–112
Егоров В. Б. Особенности многоядерности и многопоточности в сетевых процессорах	1	82–92
Егорова А. Ю., Зацман И. М., Косарик В. В., Нуриев В. А. Нестабильность нейронного машинного перевода	2	124–135
Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Машинный перевод: индикаторная оценка результатов обучения искусственной нейронной сети	4	124–137
Егорова А. Ю., Зацман И. М., Кружков М. Г., Нуриев В. А. Методика темпоральной оценки нестабильности машинного перевода	3	67–80
Ерофеева А. В., Захарова Т. В. Статистический анализ биосигналов при изучении зрительной рабочей памяти человека	4	50–60

Жуков Д. В. см. Борисов А. В.	
Зарядов И. С. см. Мейханаджян Л. А.	
Захаров В. Н. см. Ковалев Д. Ю	
Захаров В. Н. см. Плеханов Л. П.	
Захарова Т. В. см. Ерофеева А. В.	
Зацаринный А. А., Ионенков Ю. С. Методический подход к оценке вклада информационных систем в эффективность организационных систем	2 136–144
Зацаринный А. А., Королёв В. И., Гаранин А. И. Анализ влияния надежности оборудования ситуационных центров на информационную безопасность	1 93–107
Зацаринный А. А., Сучков А. П. Угрозы и риски реализации комплексных научно-технических программ в рамках приоритетов Стратегии научно-технологического развития России	3 97–111
Зацаринный А. А. см. Гаврилов В. Е.	
Зацаринный А. А. см. Грушо А. А.	
Зацман И. М. см. Вакуленко В. В.	
Зацман И. М. см. Егорова А. Ю.	
Зацман И. М. см. Егорова А. Ю.	
Зацман И. М. см. Егорова А. Ю.	
Иванов А. В. см. Борисов А. В.	
Ильин А. В., Ильин В. Д. Блокчейн-технологии нормализованного бюджетного обеспечения национальных проектов	4 147–158
Ильин А. В., Ильин В. Д. Нормализованный экономический механизм: цифровые технологии поливалютного рынка	1 186–197
Ильин А. В., Ильин В. Д. Персональные и корпоративные электронные банки в технологиях товарно-денежного обращения	2 187–194
Ильин А. В., Ильин В. Д. Ситуационная модель универсального договора в цифровой среде	3 154–162
Ильин В. Д. см. Ильин А. В.	
Ильин В. Д. см. Ильин А. В.	
Ильин В. Д. см. Ильин А. В.	
Ильин В. Д. см. Ильин А. В.	
Ионенков Ю. С. см. Зацаринный А. А.	
Кабанов А. С., Водолаженко А. А. Определение ценности и защита инсайдерской информации	2 31–42
Кириков И. А. см. Листопад С. В.	
Ковалев Д. Ю., Тарасов Е. А., Захаров В. Н., Филимонов Н. М. Архитектура платформы для проведения движимых гипотезами виртуальных экспериментов	2 56–67

	№	Стр.
Ковалев Д. Ю., Шанин И. А., Тириков Е. М. Мультидисциплинарные задачи нейрофизиологии для исполнения на распределенных вычислительных инфраструктурах	2	43–55
Ковалёв С. П. Проектирование информационного обеспечения цифровых двойников энергетических систем	1	66–81
Колесников А. В. см. Румовская С. Б.		
Коновалов Г. М. Проблема задания коэффициентов готовности диагностических систем в проекте ИТЭР: постановка задачи и алгоритм решения	2	89–102
Кончеков Е. М. см. Смирнов В. А.		
Королёв В. И. см. Зацаринный А. А.		
Косарик В. В. см. Егорова А. Ю.		
Кривенко М. П. Байесовская классификация серий многомерных данных	1	34–45
Кривенко М. П. Программное обеспечение исследований в области статистического анализа данных	4	4–13
Кривошеев И. А., Линник М. А. Статический способ стеганографического встраивания информации на основе LSB	3	56–66
Кружков М. Г. см. Егорова А. Ю.		
Кружков М. Г. см. Егорова А. Ю.		
Кудрявцев А. А., Шестаков О. В. Усредненная вероятность ошибки вычисления коэффициентов вейвлет–вейглет-разложения при обращении преобразования Радона	4	14–24
Ларичев В. А. см. Смирнов В. А.		
Линник М. А. см. Кривошеев И. А.		
Листопад С. В. Функциональная структура сплоченной гибридной интеллектуальной многоагентной системы	2	78–88
Листопад С. В., Кириков И. А. Метод идентификации конфликтов агентов в гибридных интеллектуальных многоагентных системах	1	56–65
Листопад С. В., Румовская С. Б. Нечеткое управление гетерогенным мышлением агентов гибридной интеллектуальной многоагентной системы	4	38–49
Максимов Г. А. см. Смирнов В. А.		
Малашенко Ю. Е., Назарова И. А. Оценки распределения потоков при предельной загрузке многопользовательской сети	3	4–13
Маньяков Ю. А., Сорокин А. И. Метод сегментации изображений на основе квадродерева	4	102–112
Медведева Е. Г., Хайров Э. М., Поляков Н. А., Гайдамака Ю. В. Применение кластеризации в задачах размещения подвижных точек доступа в воздушно– наземных беспроводных сетях	4	25–37

№ Стр.

Мейханаджян Л. А., Зарядов И. С., Милованова Т. А. Стационарные характеристики двухузловой марковской системы массового обслуживания с обобщенным обновлением	3	14–31
Милованова Т. А. см. Мейханаджян Л. А.		
Мистрюков А. В. Об эргодичности одноканальных систем массового обслуживания с дисциплиной абсолютного приоритета	3	32–38
Морозов Н. В. см. Степченков Ю. А.		
Морозов Н. В. см. Степченков Ю. А.		
Морозов Н. В. см. Степченков Ю. А.		
Назарова И. А. см. Малашенко Ю. Е.		
Никишин Д. А. Структура и особенности генерализации в контексте функционирования темпоральной базы геоданных	1	147–159
Никишин Д. А. см. Дулин С. К.		
Николаев А. В. см. Грушо А. А.		
Николаев А. В. см. Грушо А. А.		
Нуриев В. А. см. Егорова А. Ю.		
Нуриев В. А. см. Егорова А. Ю.		
Нуриев В. А. см. Егорова А. Ю.		
Нуриев В. А. см. Рычихин А. К.		
Нуриев В. А. см. Рычихин А. К.		
Писковский В. О. см. Грушо А. А.		
Писковский В. О. см. Грушо А. А.		
Плеханов Л. П., Захаров В. Н. Универсальный функциональный метод анализа больших самосинхронных схем	2	11–20
Поляков Н. А. см. Медведева Е. Г.		
Рацеев С. М., Череватенко О. И. Об алгоритмах декодирования обобщенных кодов Рида–Соломона	4	83–94
Рождественский Ю. В. см. Степченков Ю. А.		
Рождественский Ю. В. см. Степченков Ю. А.		
Румовская С. Б., Колесников А. В. Решение гомогенных диагностических задач специалистом в медицине	3	39–48
Румовская С. Б. см. Листопад С. В.		
Рычихин А. К., Нуриев В. А. Организация жизненного цикла исследовательских данных	3	81–96
Рычихин А. К., Нуриев В. А. Планирование жизненного цикла информационных ресурсов в (пост)грантовый период	1	135–146
Сенчило В. В. см. Грушо А. А.		
Сенчило В. В. см. Грушо А. А.		
Сенчило В. В. см. Грушо А. А.		
Синицын В. И. см. Синицын И. Н.		

	№	Стр.
Синицын И. Н., Синицын В. И. Фильтрация и экстраполяция процессов в миграционно-популяционных стохастических системах	1	4–19
Скворцов Н. А. Спецификации предметной области для решения задач с интенсивным использованием данных в нейрофизиологии	1	20–33
Скворцова Н. Н. см. Смирнов В. А.		
Смирнов В. А., Скворцова Н. Н., Кончеков Е. М., Ларичев В. А., Максимов Г. А. Программный комплекс для позиционирования акустических донных систем	4	113–123
Сорокин А. И. см. Маньяков Ю. А.		
Степченков Д. Ю. см. Степченков Ю. А.		
Степченков Д. Ю. см. Степченков Ю. А.		
Степченков Д. Ю. см. Степченков Ю. А.		
Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченков Д. Ю., Дьяченко Д. Ю. Устойчивость самосинхронного конвейера к логическим сбоям в комбинационной части	3	49–55
Степченков Ю. А., Дьяченко Ю. Г., Рождественский Ю. В., Морозов Н. В., Степченков Д. Ю., Дьяченко Д. Ю. Устойчивость самосинхронных комбинационных схем к кратковременным логическим сбоям	2	4–10
Степченков Ю. А., Морозов Н. В., Дьяченко Ю. Г., Хилько Д. В., Степченков Д. Ю. Развитие гибридной многоядерной рекуррентной архитектуры на ПЛИС	4	95–101
Сучков А. П. Алгоритмы аналитической обработки данных системы управления научными сервисами для планирования и выполнения междисциплинарных научных исследований	3	112–122
Сучков А. П. Состав, структура и методы анализа данных систем управления научными сервисами	2	163–176
Сучков А. П. см. Босов А. В.		
Сучков А. П. см. Босов А. В.		
Сучков А. П. см. Зацаринный А. А.		
Сушко Д. В. см. Борисов А. В.		
Тарасов Е. А. см. Ковалев Д. Ю.		
Терёхина И. Ю., Грушо А. А., Тимонина Е. Е., Шоргин С. Я. Построение моделей процесса с помощью простых сетей Петри	4	61–75
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		
Тимонина Е. Е. см. Грушо А. А.		

№ Стр.

- Тимонина Е. Е.** см. Терёхина И. Ю.
- Тириков Е. М.** см. Ковалев Д. Ю
- Филимонов Н. М.** см. Ковалев Д. Ю
- Хайров Э. М.** см. Медведева Е. Г.
- Хилько Д. В.** см. Степченков Ю. А.
- Череватенко О. И.** см. Рацеев С. М.
- Читалов Д. И.** Разработка модуля для измельчения ячеек расчетных сеток в нескольких направлениях и его интеграция в GUI для программной среды OpenFOAM 3 133–144
- Читалов Д. И.** Разработка приложения с графическим интерфейсом пользователя для проведения стресс-анализа на базе решателя solidDisplacementFoam программной среды OpenFOAM 1 171–185
- Шанин И. А.** см. Ковалев Д. Ю.
- Шестаков О. В.** см. Кудрявцев А. А.
- Шоргин С. Я.** см. Терёхина И. Ю.

2020 AUTHOR INDEX

	No.	Page
Adamovich I. M. and Volkov O. I. Automated estimation of the trustworthiness of concrete historical facts	1	160–170
Adamovich I. M. and Volkov O. I. Automated search for contradictions in concrete-historical information	3	145–153
Adamovich I. M. and Volkov O. I. Mechanism of temporal comparison of concrete-historical facts	4	138–146
Adamovich I. M. and Volkov O. I. Model of automated estimating the trustworthiness of concrete-historical facts system functioning	2	177–186
Biryukova T. K. Parabolic integrodifferential splines as activation functions to increase the efficiency of information processing by neural networks	4	159–167
Borisov A. V., Bosov A. V., Zhukov D. V., Ivanov A. V., and Sushko D. V. Information aspects of security in transport: Ontology of the subject area, models, and cases	1	126–134
Bosov A. V. and Suchkov A. P. Information and analytical situational center for strategic planning in the field of national security: Factors of influence	1	115–125
Bosov A. V. and Suchkov A. P. Information and analytical situational center for strategic planning in the field of national security: Perspective architectural solutions	2	145–162
Bosov A. V. see Borisov A. V.		
Cherevatenko O. I. see Ratseev S. M.		
Chitalov D. I. Development of a module for grinding meshes cells in several directions and its integration into the GUI for the OpenFOAM software environment	3	133–144
Chitalov D. I. Development of an application with a graphical user interface for stress analysis based on the solidDisplacementFoam solver of the OpenFOAM software environment	1	171–185
Diachenko D. Yu. see Stepchenkov Yu. A.		
Diachenko D. Yu. see Stepchenkov Yu. A.		
Diachenko Yu. G. see Stepchenkov Yu. A.		
Diachenko Yu. G. see Stepchenkov Yu. A.		
Diachenko Yu. G. see Stepchenkov Yu. A.		
Dulin S. K. and Nikishin D. A. Approaches to the integration of the application conceptual schemas in the unified geoontology	2	68–77
Durnovo A. A. see Vakulenko V. V.		

	No.	Page
Egorov V. B. Multicore and multithreading peculiarities in network processors	1	82–92
Egorov V. B. Some issues of software-defined data centers	2	103–112
Egorova A. Yu., Zatsman I. M., Kosarik V. V., and Nuri-ev V. A. Instability of neural machine translation	2	124–135
Egorova A. Yu., Zatsman I. M., Kruzhkov M. G., and Nuri-ev V. A. Machine translation: Indicator-based evaluation of training progress in neural processing	4	124–137
Egorova A. Yu., Zatsman I. M., Kruzhkov M. G., and Nuri-ev V. A. The technique allowing for temporal estimation of machine translation instability	3	67–80
Erofeeva A. V. and Zakharova T. V. Biosignal statistical analysis in the study of human visual working memory	4	50–60
Filimonov N. M. see Kovalev D. Y.		
Gaidamaka Yu. V. see Medvedeva E. G.		
Garanin A. I. see Zatsarinny A. A.		
Gavrilov V. E. and Zatsarinny A. A. Actual information security problems in coronavirus pandemic conditions	3	123–132
Goncharov A. A. see Vakulenko V. V.		
Grinchenko S. N. Basic information technologies and “industrial revolutions”: What is their relationship?	3	163–171
Grusho A. A., Grusho N. A., Nikolaev A. V., Piskovski V. O., Senchilo V. V., and Timonina E. E. Generation of threats in heterogeneous systems of digital economy	1	46–55
Grusho A. A., Grusho N. A., Senchilo V. V., and Timonina E. E. Prevention of mass failure of low-security mobile systems of information support for digital economy	2	21–30
Grusho A. A., Nikolaev A. V., Piskovski V. O., Senchilo V. V., and Timonina E. E. End-to-end information security of private cloud computing	4	76–82
Grusho A. A., Zatsarinny A. A., and Timonina E. E. Description of digital economy dynamics using an electronic ledger	1	108–114
Grusho A. A. see Teryokhina I. Yu.		
Grusho N. A. see Grusho A. A.		
Grusho N. A. see Grusho A. A.		
Ilyin A. V. and Ilyin V. D. Blockchain technologies for normalized budget support of national projects	4	147–158
Ilyin A. V. and Ilyin V. D. Personal and corporate e-banks in commodity-money circulation technologies	2	187–194
Ilyin A. V. and Ilyin V. D. Situational model of universal agreement in a digital environment	3	154–162

	No.	Page
Ilyin A. V. and Ilyin V. D. The normalized economic mechanism: Digital technologies of the multicurrency market	1	186–197
Ilyin V. D. see Ilyin A. V.		
Ilyin V. D. see Ilyin A. V.		
Ilyin V. D. see Ilyin A. V.		
Ilyin V. D. see Ilyin A. V.		
Ionenkov Yu. S. see Zatsarinny A. A.		
Ivanov A. V. see Borisov A. V.		
Kabanov A. S. and Vodolazhenko A. A. Valuation and protection of insider information	2	31–42
Khayrov E. M. see Medvedeva E. G.		
Khilko D. V. see Stepchenkov Yu. A.		
Kirikov I. A. see Listopad S. V.		
Kolesnikov A. V. see Rumovskaya S. B.		
Konchekov E. M. see Smirnov V. A.		
Konovalov G. M. The problem of setting availability targets for ITER diagnostics: Problem statement and solution algorithm	2	89–102
Korolev V. I. see Zatsarinny A. A.		
Kosarik V. V. see Egorova A. Yu.		
Kovalev D. Y., Shanin I. A., and Tirikov E. M. Multidisciplinary neuroinformatics problems for execution in distributed computing infrastructures	2	43–55
Kovalev D. Y., Tarasov E. A., Zakharov V. N., and Filimonov N. M. Architecture of the platform for managing hypotheses-driven virtual experiments	2	56–67
Kovalyov S. P. Information architecture of the power system digital twin	1	66–81
Krivenko M. P. Bayesian classification of serial multivariate data	1	34–45
Krivenko M. P. Software of research in statistical data analysis	4	4–13
Krivosheev I. A. and Linnik M. A. Static method for steganographic embedding based on least significant bit	3	56–66
Kruzhkov M. G. see Egorova A. Yu.		
Kruzhkov M. G. see Egorova A. Yu.		
Kudryavtsev A. A. and Shestakov O. V. Average probability of error in calculating wavelet–vaguelette coefficients while inverting the Radon transform	4	14–24
Larichev V. A. see Smirnov V. A.		
Linnik M. A. see Krivosheev I. A.		
Listopad S. V. Functional structure of the cohesive hybrid intelligent multiagent system	2	78–88
Listopad S. V. and Kirikov I. A. Agent conflict identification method in hybrid intelligent multiagent systems	1	56–65

- Listopad S. V. and Rumovskaya S. B.** Fuzzy control of heterogeneous thinking of the hybrid intelligent multiagent system's agents 4 38–49
- Malashenko Yu. E. and Nazarova I. A.** Flows distribution estimations at maximum load of multiuser network 3 4–13
- Maniakov Yu. A. and Sorokin A. I.** Quadtree based color image segmentation method 4 102–112
- Maximov G. A.** see Smirnov V. A.
- Medvedeva E. G., Khayrov E. M., Polyakov N. A., and Gaidamaka Yu. V.** Application of clustering in deployment of mobile access points in air-ground wireless networks 4 25–37
- Meykhanadzhyan L. A., Zaryadov I. S., and Milovanova T. A.** Stationary characteristics of the two-node Markovian tandem queueing system with general renovation 3 14–31
- Milovanova T. A.** see Meykhanadzhyan L. A.
- Mistryukov A. V.** Ergodicity of single server queues with preemptive priority 3 32–38
- Morozov N. V.** see Stepchenkov Yu. A.
- Morozov N. V.** see Stepchenkov Yu. A.
- Morozov N. V.** see Stepchenkov Yu. A.
- Nazarova I. A.** see Malashenko Yu. E.
- Nikishin D. A.** Structure and features of generalization in the context of functioning of a temporal geodatabase 1 147–159
- Nikishin D. A.** see Dulin S. K.
- Nikolaev A. V.** see Grusho A. A.
- Nikolaev A. V.** see Grusho A. A.
- Nuriev V. A. and Rychikhin A. K.** Data life cycle planning in the (post)grant period 1 135–146
- Nuriev V. A.** see Egorova A. Yu.
- Nuriev V. A.** see Egorova A. Yu.
- Nuriev V. A.** see Egorova A. Yu.
- Nuriev V. A.** see Rychikhin A. K.
- Piskovski V. O.** see Grusho A. A.
- Piskovski V. O.** see Grusho A. A.
- Plekhanov L. P. and Zakharov V. N.** Universal functional method for analyzing large self-timed circuits 2 11–20
- Polyakov N. A.** see Medvedeva E. G.
- Ratseev S. M. and Cherevatenko O. I.** On decoding algorithms for generalized Reed–Solomon codes 4 83–94
- Rogdestvenski Yu. V.** see Stepchenkov Yu. A.
- Rogdestvenski Yu. V.** see Stepchenkov Yu. A.
- Rumovskaya S. B. and Kolesnikov A. V.** Solving of diagnostic tasks by a doctor 3 39–48

No.	Page
Rumovskaya S. B. see Listopad S. V.	
Rychikhin A. K. and Nuriev V. A. Organization of the research data life cycle	3 81–96
Rychikhin A. K. see Nuriev V. A.	
Senchilo V. V. see Grusho A. A.	
Senchilo V. V. see Grusho A. A.	
Senchilo V. V. see Grusho A. A.	
Shanin I. A. see Kovalev D. Y.	
Shestakov O. V. see Kudryavtsev A. A.	
Shorgin S. Ya. see Teryokhina I. Yu.	
Sinitsyn I. N. and Sinitsyn V. I. Filtering and extrapolation in migrational-populational stochastic systems	1 4–19
Sinitsyn V. I. see Sinitsyn I. N.	
Skvortsov N. A. Domain specifications for data-intensive problem solving in neurophysiology	1 20–33
Skvortsova N. N. see Smirnov V. A.	
Smirnov V. A., Skvortsova N. N., Konchekov E. M., Larichev V. A., and Maximov G. A. Software package for positioning of acoustic bottom systems	4 113–123
Sorokin A. I. see Maniakov Yu. A.	
Stepchenkov D. Yu. see Stepchenkov Yu. A.	
Stepchenkov D. Yu. see Stepchenkov Yu. A.	
Stepchenkov D. Yu. see Stepchenkov Yu. A.	
Stepchenkov Yu. A., Diachenko Yu. G., Rogdestvenski Yu. V., Morozov N. V., Stepchenkov D. Yu., and Diachenko D. Yu. Self-timed combinational circuit tolerance to short-term soft errors	2 4–10
Stepchenkov Yu. A., Diachenko Yu. G., Rogdestvenski Yu. V., Morozov N. V., Stepchenkov D. Yu., and Diachenko D. Yu. Self-timed pipeline immunity to soft errors in its combinational part	3 49–55
Stepchenkov Yu. A., Morozov N. V., Diachenko Yu. G., Khilko D. V., and Stepchenkov D. Yu. Multicore hybrid recurrent architecture expansion on FPGA	4 95–101
Suchkov A. P. Algorithms for an analytical data processing system of management of scientific services for planning and execution of interdisciplinary research	3 112–122
Suchkov A. P. Content, structure, and data analysis methods of scientific services management system	2 163–176
Suchkov A. P. see Bosov A. V	
Suchkov A. P. see Bosov A. V.	
Suchkov A. P. see Zatsarinny A. A.	

	No.	Page
Sushko D. V. see Borisov A. V.		
Tarasov E. A. see Kovalev D. Y.		
Teryokhina I. Yu., Grusho A. A., Timonina E. E., and Shor- gin S. Ya. Constructing process models represented by simple Petri nets	4	61–75
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Grusho A. A.		
Timonina E. E. see Teryokhina I. Yu.		
Tirikov E. M. see Kovalev D. Y.		
Vakulenko V. V., Goncharov A. A., Durnovo A. A., and Zats- man I. M. Tasks of the phraseological dictionary database and stages of its design	2	113–123
Vodolazhenko A. A. see Kabanov A. S.		
Volkov O. I. see Adamovich I. M.		
Volkov O. I. see Adamovich I. M.		
Volkov O. I. see Adamovich I. M.		
Volkov O. I. see Adamovich I. M.		
Zakharov V. N. see Kovalev D. Y.		
Zakharov V. N. see Plekhanov L. P.		
Zakharova T. V. see Erofeeva A. V.		
Zaryadov I. S. see Meykhanadzhyan L. A.		
Zatsarinny A. A. and Ionenkov Yu. S. Methodological approach to assessing the contribution of information systems to the effectiveness of organizational systems	2	136–144
Zatsarinny A. A., Korolev V. I., and Garanin A. I. Analysis of the impact of reliability indicators equipment of situation centers on information security	1	93–107
Zatsarinny A. A. and Suchkov A. P. Threats and risks of imple- menting complex scientific and technical programs within the priorities of the Russian science and technology development strategy	3	97–111
Zatsarinny A. A. see Gavrilov V. E.		
Zatsarinny A. A. see Grusho A. A.		
Zatsman I. M. see Egorova A. Yu.		
Zatsman I. M. see Egorova A. Yu.		
Zatsman I. M. see Egorova A. Yu.		
Zatsman I. M. see Vakulenko V. V.		
Zhukov D. V. see Borisov A. V.		

Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация предоставленной автором(ами) рукописи не должна нарушать положений глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редактуре должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L^AT_EX (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:

- название статьи;
- Ф.И.О. авторов, на английском можно только имя и фамилию;
- место работы, с указанием города и страны и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp и
http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp;
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.).

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.

Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

Описание статьи из журнала:

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Описание статьи из электронного журнала:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Описание материалов конференций:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Особенности проектирования разработки месторождений с применением гидравлического трещинообразования [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povышeniya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

Описание книги (монографии, сборники):

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Описание переводной книги (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

Timoshenko S. P., Young D. H., Weaver W. Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (Timoshenko S. P., Young D. H., Weaver W. Vibration problems in engineering. — 4th ed. — New York, NY, USA: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York, NY: Wiley. 521 p.

Описание неопубликованного документа:

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

Описание интернет-ресурса:

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

Описание диссертации или автореферата диссертации:

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. Moscow. D.Sc. Diss. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informacionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

Описание ГОСТа:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publs. 10 p.

Описание патента:

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.
11. При отправке файлов по электронной почте просим придерживаться следующих правил:
 - указывать в поле subject (тема) название журнала и фамилию автора;
 - использовать attach (присоединение);
 - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

Адрес редакции журнала «Системы и средства информатики»:

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: rust@ipiran.ru (Стригина Светлана Николаевна)

<http://www.ipiran.ru/journal/collected>

Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

Agreement on the transfer of rights to publish:

“We, the undersigned authors of the manuscript “. . .,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.”

Author(s) signature(s): (name(s), address(es), date).”

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.
6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L^AT_EX (.tex) formats, and additionally in the .pdf format. All documents

may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.

7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font —Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:

<http://www.ipiran.ru/publication/collected/template.doc>

8. Articles should enclose data both in **Russian and English**:

- title;
- author's name and surname;
- affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
- data on authors according to the format (see site):
http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp and
http://www.ipiran.ru/journal/collected/2019_29_03_eng/authors.asp;
- abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
- Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.
Important! Keywords must not be sentences.
- Acknowledgments.

9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.net/ru/bgn/>).

Please take into account the following examples of Russian references appearance:

Article in journal:

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Russ. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

Journal article in electronic format:

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

Conference proceedings:

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursosberegayushchie tekhnologii nedropol'zovaniya i povышeniya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

Books and other monographs:

Lindorf, L. S., and L. G. Mamikonants, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

Dissertation and Thesis:

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 23 p.

State standards and patents:

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoj samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
 - the journal title and author's name in the "Subject" field;
 - an article and additional materials have to be attached using the "attach" function;
 - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

Editorial Board address:

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499)135 86 92, Fax: +7 (495)930 45 05

e-mail: rust@ipiran.ru (to Svetlana Strigina)

http://www.ipiran.ru/english/journal_systems.asp

SYSTEMS AND MEANS OF INFORMATICS (СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ)

SCIENTIFIC JOURNAL

Volume 30 No.4 Year 2020

Editor-in-Chief and Chair of Editorial Council
Academician I. A. Sokolov

I N T H I S I S S U E:

SOFTWARE OF RESEARCH IN STATISTICAL DATA ANALYSIS <i>M. P. Krivenko</i>	4
AVERAGE PROBABILITY OF ERROR IN CALCULATING WAVELET–VAGUELETTE COEFFICIENTS WHILE INVERTING THE RADON TRANSFORM <i>A. A. Kudryavtsev and O. V. Shestakov</i>	14
APPLICATION OF CLUSTERING IN DEPLOYMENT OF MOBILE ACCESS POINTS IN AIR–GROUND WIRELESS NETWORKS <i>E. G. Medvedeva, E. M. Khayrov, N. A. Polyakov, and Yu. V. Gaidamaka</i>	25
FUZZY CONTROL OF HETEROGENEOUS THINKING OF THE HYBRID INTELLIGENT MULTIAGENT SYSTEM'S AGENTS <i>S. V. Listopad and S. B. Rumovskaya</i>	38
BIOSIGNAL STATISTICAL ANALYSIS IN THE STUDY OF HUMAN VISUAL WORKING MEMORY <i>A. V. Erofeeva and T. V. Zakharova</i>	50
CONSTRUCTING PROCESS MODELS REPRESENTED BY SIMPLE PETRI NETS <i>I. Yu. Teryokhina, A. A. Grusho, E. E. Timonina, and S. Ya. Shorgin</i>	61
END-TO-END INFORMATION SECURITY OF PRIVATE CLOUD COMPUTING <i>A. A. Grusho, A. V. Nikolaev, V. O. Piskovski, V. V. Senchilo, and E. E. Timonina</i>	76
ON DECODING ALGORITHMS FOR GENERALIZED REED–SOLOMON CODES <i>S. M. Ratseev and O. I. Cherevatenko</i>	83
MULTICORE HYBRID RECURRENT ARCHITECTURE EXPANSION ON FPGA <i>Yu. A. Stepchenkov, N. V. Morozov, Yu. G. Diachenko, D. V. Khilko, and D. Yu. Stepchenkov</i>	95